Received: 30 September 2022 | Accepted: 15 March 2023

DOI: 10.1111/bjet.13317

ORIGINAL ARTICLE

British Journal of Educational Technology BERA

Pedagogical features of interactive apps for effective learning of foundational skills

Bethany Huntington[1] | **James Goulding**[2] | **Nicola J. Pitchford**[1]

[1]School of Psychology, University of Nottingham, Nottingham, UK

[2]N/LAB, University of Nottingham, Nottingham, UK

Correspondence

Bethany Huntington, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, UK.
Email: bethany.huntington@nottingham.ac.uk

Funding information

Economic and Social Research Council, Grant/Award Number: ES/J500100/1

Interactive apps are commonly used to support the acquisition of foundational skills. Yet little is known about how pedagogical features of such apps affect learning outcomes, attainment and motivation—particularly when deployed in lower-income contexts, where educational gains are most needed. In this study, we analyse which app features are most effective in supporting the acquisition of foundational literacy and numeracy skills. We compare five apps developed for the Global Learning XPRIZE and deployed to 2041 out-of-school children in 172 remote Tanzanian villages. A total of 41 non-expert participants each provided 165 comparative judgements of the five apps from the competition, across 15 pedagogical features. Analysis and modelling of these 6765 comparisons indicate that the apps created by the joint winners of the XPRIZE, who produced the greatest learning outcomes over the 15-month field trial, shared six pedagogical features—autonomous learning, motor skills, task structure, engagement, language demand and personalisation. Results demonstrate that this combination of features is effective at supporting learning of foundational skills and has a positive impact on educational outcomes. To maximise learning potential in environments with both limited resources and deployment opportunities, developers should focus attention on this combination of features, especially for out-of-school children in low- and middle-income countries.

KEYWORDS
accessibility, app features, children, comparative judgement, educational technology, learning apps, remote learning

Practitioner notes

What is already known about this topic

- Interactive apps are becoming common to support foundational learning for children both in and out of school settings.
- The Global Learning XPRIZE competition demonstrates that learning apps can facilitate learning improvements in out-of-school children living in sub-Saharan Africa.
- To understand which app features are most important in supporting learning in these contexts, we need to establish which pedagogical features were shared by the winning apps.

What this paper adds

- Effective learning of foundational skills can be achieved with a range of pedagogical features.
- To maximise learning, apps should focus on combining elements of autonomous learning, motor skills, task structure, engagement, language demand and personalisation.
- Free Play is not a key pedagogical feature to facilitate learning within this context.

Implications for practice and/or policy

- When developing learning apps with primary-aged, out-of-school children in low-income contexts, app developers should try to incorporate the six key features associated with improving learning outcomes.
- Governments, school leaders and parents should use these findings to inform their decisions when choosing an appropriate learning app for children.

INTRODUCTION

Interactive apps are becoming increasingly common to support the acquisition of foundational skills (ie, basic literacy and numeracy) in primary school classrooms and home learning environments around the world (Alemán de la Garza et al., 2019). A growing evidence base suggests that using curriculum-based apps can provide high-quality education to children globally, supporting positive outcomes in literacy development, maths, science, problem-solving and self-efficacy (Bettinger et al., 2020; Herodotou, 2018). Literacy and numeracy interventions in both developing and developed countries, implemented both at school and in a home environment, have shown significant improvements in attainment, learning outcomes and motivation for children (Major & Francis, 2020; Stubbé et al., 2016).

There was an exponential increase in mobile learning apps across a 3-month period in 2019–2020, with over 900 million learning apps downloaded worldwide (Statista, 2021). As the COVID-19 pandemic took a grip on educational provision globally, resulting in school closures that left approximately 1.2 billion children unable to attend school worldwide (Forbes, 2021), dependence on tablet-based apps increased further, and emphasised the need for alternative, effective education provision that is feasible when schools are closed, or children are not able to attend school (Azevedo et al., 2021).

When successful, interactive apps can offer exciting and effective learning environments that foster child-centred learning and provide a stark contrast to traditional teacher-centred learning styles (Papadakis & Kalogiannakis, 2017; Ting, 2015). Learning apps currently on the market vary vastly in structure, content and quality, so unless they are evaluated scientifically, researchers, educators and parents cannot establish how effective apps are in supporting learning (Kolak et al., 2021).

Despite an accelerated uptake of mobile learning apps, relatively little is known about which app features support positive learning outcomes (Kim et al., 2021). When apps are used outside of the school setting, without the support of a teacher or caregiver to promote the acquisition of foundational skills, it is vital to decipher app features that are effective in assisting children's learning.

Over the past decade, several app features have been proposed to support learning, derived from evaluation tools created to assess the design and potential of learning apps (eg, Kolak et al., 2021; Outhwaite et al., 2022; Papadakis et al., 2017). Well-chosen interactive features embedded into mobile apps can help facilitate child-paced, inclusive learning environments. Table 1 describes 15 key app features that have been attributed by different researchers to support learning and contribute to the educational value of an app. While not an exhaustive or partially systematic review of each paper, these features reflect a categorisation of attributes commonly identified in the literature as supporting learning, until saturation for the purpose of this study was reached. Only pedagogical app features are reported; technical characteristics of tablets that may impact learning, such as screen size, are not considered.

Although these app features have been championed by educational researchers, it is not possible to know how important these features are in supporting learning until direct comparisons of different apps are made, that have been deployed in the same context, with known learning outcomes. This is crucial, as some of the app features are in direct contradiction with each other. For example, direct instruction takes a prescribed, structured approach to learning, whereas Free Play is more child-centred and provides children with the autonomy to explore in an unstructured manner (Hirsh-Pasek et al., 2015; Outhwaite, Faulder, et al., 2019). Playful learning that combines direct instruction with Free Play has been demonstrated in school environments and with app-based learning and could be highly effective in promoting socio-emotional and cognitive development in primary-aged children as both are shown to have unique benefits (Hirsh-Pasek et al., 2015; Toub et al., 2016). Apps that embody both features afford instruction through gradual release of responsibility, whereby cognitive work slowly and gradually shifts from tutor-led direction to student-led discovery (Fisher & Frey, 2021; Northrop & Killeen, 2013). Hence, it is important to identify the combinations of features embedded within mobile apps that have been shown to be effective at promoting learning outcomes.

To understand how individual or combinations of app features support learning, app features need to be linked directly to learning outcomes. A first step towards this goal was made by Outhwaite et al. (2022) in a systematic review and content analysis of maths apps targeted at the first 3 years of compulsory schooling that had been previously evaluated in the literature, as well as the Top 25 learning apps on the App Stores. They conducted a qualitative comparative analysis (QCA), which identified specific app design features, and combination of features, that were shown to be sufficient to support children's learning of maths with educational apps. Out of 50 studies included in their systematic review, only 8 apps met the criteria required for the QCA of having data reported in sufficient detail to enable within-subject effect sizes (Cohen's d) to be calculated. Results of the QCA revealed that features of the maths apps that promoted programmatic levelling, such as scaffolding

TABLE 1 App features that have been attributed to support children's learning.

App feature	Definition	Cited by
Active learning	Is the app 'minds on'? Does it require thinking or intellectual effort (rather than just cause-and-effect interactions or guessing)?	Hirsh-Pasek et al. (2015) Plump and LaRosa (2017)
Engagement in the learning process	Do the features engage you in app activities? Or do they distract you? Are the visual and sound effects excessive? Are there any disruptive ads?	Hirsh-Pasek et al. (2015) Plump and LaRosa (2017) Lee and Cherner (2015)
Meaningful learning	Is the content meaningful and relevant to children's everyday experiences? Is it taught in a manner that can be contextualised within existing knowledge?	Hirsh-Pasek et al. (2015) Kolak et al. (2021)
Social interaction	To what extent can the children interact meaningfully with (a) characters through the app interface and (b) caregivers around the app?	Hirsh-Pasek et al. (2015) Kolak et al. (2021) Lee and Cherner (2015) Outhwaite et al. (2022)
Accessibility—language demand	Is the language used simple enough to be accessible to children with Special Educational Needs and Disabilities (SEND) or children with lower language proficiency?	Kolak et al. (2021) Gulliford et al. (2021)
Accessibility—motor skills	Would the app be suitable for children with lower motor skills? Would it be good for children who usually struggle with traditional paper and pencil skills?	Allison (2019) Gulliford et al. (2021) Pitchford et al. (2016) Pitchford et al. (2018)
Accessibility—autonomous learning	Is the app easy to navigate independently and signposted appropriately for someone with limited tech experience? Is the avatar (if there is one) helpful in guiding the child? If no external caregiver, do you think a child could navigate the app with some trial and error?	Papadakis et al. (2017) Lan (2018)
Task structure	Are the tasks structured in a way that makes sense? Is the child directed to which tasks they should complete next? Is there opportunity to reinforce previously learnt skills/knowledge?	Gulliford et al. (2021) Outhwaite et al. (2022) Callaghan and Reich (2018)
Task processes—feedback	Does the app provide both positive and negative feedback? When negative, does the error signal come with linked instructional feedback (to help them understand what they are doing wrong)? Is the feedback encouraging and potentially exciting for children (eg, getting a prize for completing a task)?	Kolak et al. (2021) Gulliford et al. (2021) Outhwaite et al. (2022) Benton et al. (2021)

TABLE 1 (Continued)

App feature	Definition	Cited by
Curriculum links	Do you think the tasks are close to what would be taught in schools on the curriculum? Do you think they are necessary topics/things children should know? Do you think the apps could be used to reinforce things taught in class? Do you think the apps could be used as an assessment tool for curriculum modules?	Gulliford et al. (2021) Richards (2015)
Gamification	Are there games within the app? Are these fun and engaging? Is there a reward for 'winning' the game or being successful, or different levels to complete?	Putra et al. (2018) Lee and Loo (2021) Al-Azawi et al. (2016)
Personalisation/personalised levelling	Is the app personalised to the child? Does it provide tasks that are specific to the child's level of learning? Do these tasks adjust accordingly as the child progresses? Or is it a 'one task fits all' approach?	Outhwaite et al. (2022) Lee and Cherner (2015) Benton et al. (2021) Vanbecelaere et al. (2020) Vanbecelaere et al. (2021)
Retrieval-based learning	Are there quizzes/tasks to test what children have previously learnt in the app?	Pitchford (2015) Grimaldi and Karpicke (2014)
Direct instruction	Does a teacher/avatar give explicit instructions for how to do something? Do they demonstrate this or is there an example shown, using a demo, written example or video? Do they use the 'show, try, test' method?	Outhwaite, Faulder, et al. (2019) Toub et al. (2016) Chodura et al. (2015)
Free play	Is the play unstructured? Are children free to explore as they want to? Are the activities child-centred? Do any of the tasks allow children to use their imagination?	Toub et al. (2016) Hirsh-Pasek et al. (2015)

and personalisation, as well as explanatory and motivational feedback, maximised children's learning outcomes. Outhwaite et al. (2022) called for learning apps to be evaluated in different settings, including the home environment, to enhance understanding of how they might address educational challenges faced in different contexts.

In this paper, we investigate how educational apps might address the global learning crisis faced by 244 million children worldwide, especially those residing in sub-Saharan Africa, who do not attend school (UNESCO, 2022). We capitalise on a unique randomised control trial (RCT) that directly compared five learning apps developed by different teams from around the world, to support learning of foundational literacy and numeracy skills in children residing in remote villages in Tanga, Tanzania. Learners were of a similar age (7–11 years) and ability, as they were all out of school, and at the start of the study 74% could not read a single word of Swahili (XPRIZE, 2019). By exploring features of apps designed to support the acquisition of foundational literacy and numeracy skills, that were trialled within the context of the same RCT, with out-of-school children aged 7–11 years in a low-income country, this study adds to previous app evaluation frameworks that have focused on learning in one domain (eg, maths) with early years or preschool children (Kolak et al., 2021; Outhwaite et al., 2022, respectively).

The Global Learning XPRIZE

To address the global learning crisis, the XPRIZE Foundation launched a 'Global Learning' competition in partnership with the World Food Programme and UNESCO that challenged multi-disciplinary teams across the world to develop open-source, scalable learning software empowering children to teach themselves basic literacy, numeracy and writing skills (XPRIZE, 2019). The challenge was to develop a tablet-based, scalable, digital technology solution for marginalised out-of-school children. XPRIZE placed no requirements or constraints on app structure or content, allowing the teams to be independent in their design.

Five finalist teams field tested their app with 2041 illiterate children, aged 7–11 years, from across 172 remote villages in Tanzania over a 15-month period (XPRIZE, 2019). Each village received one app on their tablets, and this was statistically balanced to keep the competition as fairly distributed as possible. Teams were made aware that there was no formal learning support in place for children throughout this field test. Learning was independent and there were village 'Mamas' hired to help the children with any technical issues, but not with any content assistance (Huntington et al., 2021).

An independent company, RTI International, assessed the impact of each app provided by the five finalists teams. Using standardised assessments of literacy and numeracy, individual children were assessed pre- and post-intervention. While all teams demonstrated significant core improvements in literacy, maths and writing skills over the duration of the 15-month field trial, there were two joint winning apps of the GLXP—onebillion and Kitkit School—as children who received instruction with these two apps achieved the greatest overall proficiency gains compared with children who used an app produced by the other three finalist teams (XPRIZE, 2019).

Current study

As there were clear winners and objective measures of learning gains from the GLXP, it is possible to directly compare embedded features of the finalist apps and draw inferences about which features are most effective at supporting learning for out-of-school children. To do this, we conducted a comparative judgement task, in which 41 naïve participants were asked to compare the five finalist apps on the 15 app features listed in Table 1 that have been proposed to support foundational learning. Comparative judgement was chosen over other methods to avoid potential inconsistencies that can occur in use of absolute judgement-based questioning (Kalton & Schuman, 1982) and biases that can occur (eg, acquiescence bias) when using corresponding tools such as Likert Scales (eg, Kim et al., 2021). Use of pairwise comparison allowed the five finalist apps and 15 pedagogical features to be clearly and consistently ranked both in a timely fashion and with no potential for ties (Marshall et al., 2020). Dichotomous ratings assessing if a feature was present within the apps (as in Outhwaite et al., 2022; Papadakis et al., 2018) were not used given their own potential to produce ceiling and floor effects (ie, when a feature is present or absent in all apps). In contrast, the use of a comparative judgement task involved participants selecting an app that embodied a certain feature to a greater extent than another app. As such, relative judgements were able to order the extent to which different learning features were embodied within the five finalist apps.

This study mirrors previous work that has evaluated apps using a framework of features, but, for the first time, the learning gains achieved from the GLXP allow us to identify which features are most successful in facilitating children's learning of foundational skills in low-income community settings, by identifying key features of the two winning apps. To achieve this, the following questions were investigated:

1. Which features characterise each of the five finalist learning apps used in the GLXP?
2. Do the five finalist learning apps of the GLXP, that have been shown to support positive learning outcomes with out-of-school children within the same RCT, share features?

METHOD

Design

A within-participants experimental design was adopted in which all participants completed 165 trials (150 experimental trials and 15 repeated trials to assess consistency of response) of a two-alternative, forced-choice comparative judgement task (a comparison method as outlined in Pollitt, 2012). Each trial required participants to judge which of the two apps they were presented with was strongest on a particular feature (as listed in Table 1), such as *direct instruction* or availability of *autonomous learning*. In this manner, each of the five finalist apps was iteratively compared against each of the others. This gave rise to a 'comparative judgement' score per app, which reflected the number of times the app was chosen over another in reference to a particular feature, with scores ranging between zero (never favoured) and four (always favoured) per app for each of the 15 features investigated. The order of comparative judgement trials was randomised across participants.

Prior to data collection, a pilot study was conducted with five participants, after which minor adjustments were made to refine this protocol based on researcher observations and participant feedback, such as presenting the full names of the apps throughout the task for clarity (eg, CCI-School House instead of CCI). Analyses of the pilot data showed the task to be reliable and valid. All data were collected in a computer laboratory at the University of Nottingham. Ethical approval was granted by the School of Psychology Ethics Committee. Consent was obtained from all participants in line with the British Psychological Society guidelines.

Participants

Forty-one participants took part in the comparative judgement task described above. Participant ages ranged from 18 to 38 years (M = 24.14, SD = 4.51), with 27 females, 13 males and 1 non-binary participant. All participants were residents of the United Kingdom and either currently or previously enrolled in Higher Education. Participants were recruited through the School of Psychology at the University of Nottingham using opportunity sampling via email invitations, posters, flyers and word of mouth. Each participant was provided with an inconvenience allowance of £20 for taking part in the study.

All participants were blind to the study aims and had no a priori knowledge of either the results of the GLXP competition or the five apps used, as confirmed by the researcher before each session. Rather than recruiting educational experts to this study, participants were non-experts thus mitigating risks associated with participants having different levels of background knowledge and expertise (Bramley, 2007).

Apparatus and materials

Five Google Pixel C tablets with a 10.2″ screen with a resolution of 2560 × 1800 pixels were used, one for each participant in any one session. Viewing distance was not controlled. The tablets were the same as those used during the GLXP competition with out-of-school children in Tanzania. The tablets enabled participants to access and interact with the five finalist apps before making comparative judgements. A brief description of each of the five finalist apps, installed on the tablets utilised by participants, is outlined in Table 2. The same version of the apps used in this study were those used in the GLXP competition, where possible. Participants were made aware of any minor differences in app versions before completing the comparative judgement trials but were instructed not to base judgements on these differences. Specifically, the version of onebillion used in this study included a teacher loading page and the version of Robotutor used in this study did not block access to other apps, as in the XPRIZE version.

Participants used their own headphones when exploring the apps; these could not be provided due to COVID-19 precautions that needed to be followed. Pen and paper were provided to participants so that they could make notes. Each session started with a presentation delivered by the researcher (BH) that explained the GLXP, the aims of the current study and any information about deviations of the apps used in the study compared with those used in the GLXP, as described above. The session structure was outlined with instructions for each step. The results of the GLXP were not revealed at this time to avoid participant bias. All participants were provided with a booklet containing definitions for the 15 app features, framed as questions, as listed in Table 1. Some of the feature definitions were taken from specific papers or models (eg, definitions from the Four Pillars of Learning framework; Hirsh-Pasek et al., 2015), while others were synthesised from multiple papers of app learning.

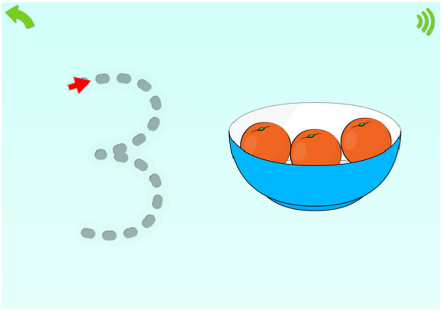
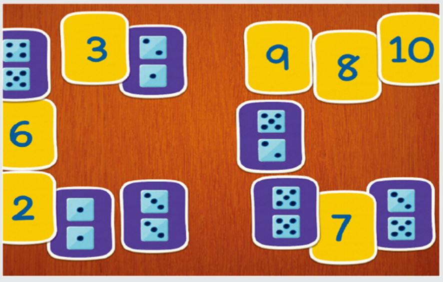
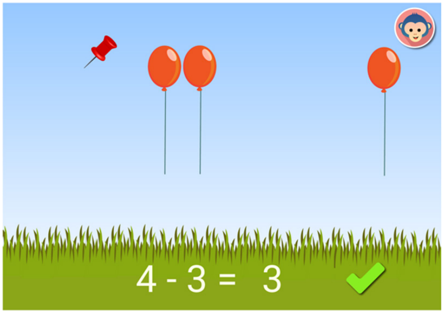
Qualtrics software, version February 2022, was used to present the comparative judgement trials and record participant responses. For each of the 15 features, there were 10 possible app pairings, resulting in 150 distinct trials. Simple syntax was used to randomise question order and avoid bias. For example, when making judgements on the app feature of 'Gamification' participants were asked: '*Which app has more gamification? onebillion or Robotutor?'* Participants were required to click on the app they felt had the most of that feature.

To maintain attention and check for active participation throughout the task, after every 30 trials, participants were required to answer an unrelated, two-alternative, forced choice question, for example, *'What season is it right now? Spring or Autumn?'*. Furthermore, for each of the 15 app features, one additional trial was repeated, but the app choice was reversed. For example, the question on Gamification given above was changed to *'Which app has more gamification? Robotutor or onebillion?* This was to determine the consistency of participants' responses throughout the task.

Procedure

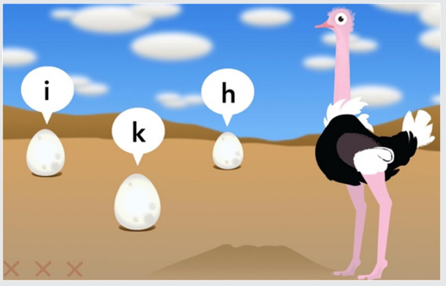
Participants contacted the researcher via email if they were interested in taking part. They were provided with an information sheet and were then invited to take part in a 2-hour session held at a time and date that was convenient for them. Each session had two to five participants, with most comprising five participants. At the beginning of each session, participants were welcomed, and the researcher presented the introductory information and task instructions, which took approximately 10–15 minutes. Participants were then asked to

TABLE 2 Brief description of the five finalist learning apps from the GLXP as described on the organisation's website.

App/website	Pedagogical description
onebillion www.onebillion.org 	The child works through an ordered set of learning units covering reading, writing and numeracy. Each is a short interactive activity to teach or provide practice for a particular concept or skill. The app provides each child with an individual journey. Regular tests enable the app to deliver the best course through the content for each individual child based on their learning level
Kitkit School www.kitkitschool.com 	A sequenced progression of core literacy skills, from letter recognition to phonics and print awareness. Sequential courses introduce new skills and reinforce previously covered concepts at more difficult levels. Learning is scaffolded to support cognitive development and independent learning and accessibility functions engage and empower the world's diverse learners
Chimple www.chimple.org 	Chimple features a series of educational games in literacy, maths and digital skills. The app is easy to use and understand. It enables children to learn, practice and improve their reading, writing and math skills in a fun and interactive way

(Continues)

TABLE 2 (Continued)

App/website	Pedagogical description
CCI www.cciny.net  	SchoolHouse (by CCI) includes a series of structured and sequential instructional lessons, as well as a platform that enables non-coders to develop engaging learning content in any language or subject area
Robotutor www.cmu.edu/scs  	Robotutor is an open-source Android tablet app that enables children ages 7–10 with little or no access to schools to learn basic reading, writing and arithmetic without adult assistance

spend 5 minutes for reading the feature definitions carefully and were given the opportunity to ask any questions for clarification.

Participants were shown how to use the tablets to access and interact with the five finalist apps. They were then asked to spend unstructured time exploring each app to familiarise themselves with the app features. During this process, it was recommended that participants make notes of anything positive, negative or different that could support them in making comparative judgements. They were instructed to spend a minimum of 25 minutes on this activity, 5 minutes per app but assured that they could spend longer familiarising themselves with the apps if they wanted to. Participants were asked to engage with a range of literacy and numeracy tasks within each app and to give both correct and incorrect answers to experience the range of feedback provided. The researcher was on hand to answer questions and troubleshoot any technical issues.

When participants felt sufficiently familiarised with the apps, the researcher checked they understood the task, answering any remaining questions they might have. Participants then performed the comparative judgement task independently, with no group discussion. Participants were instructed to work at their own pace and reminded they could refer to the apps at any point when making their judgements, allowing further, targeted exploration if needed, to enable them to make informed decisions. Participants actively engaged with the apps while completing the comparative judgement task. Once all participants had completed the task, the researcher thanked them for taking part and answered any outstanding queries. Sessions lasted between 1 hour 40 minutes and 2 hours and 10 minutes, with an average duration of 1 hour and 55 minutes. Consent forms, debriefs and withdrawal statements were provided during the session in accordance with ethical guidelines.

Data analysis

Inclusion criteria for data analysis required participants to answer 100% of the unrelated attention questions and 80% of the ‘retest’ questions accurately. All participants achieved this and consequently, no participants were excluded from the analysis.

Data were first analysed to assess the reliability of the comparative judgements: responses to the 15 repeated trials were compared with the corresponding experimental trial response for each participant. Internal consistency was determined by calculating the number of times participants gave the same response across repeated and original experimental trials, from which a group mean consistency was determined. Internal consistency for participant responses across the 15 repeated trials was high, with a group mean score of 13.83 ($SD = 1.16$; 92.2% accuracy), demonstrating comparative judgements were reliable. Fleiss' kappa was then conducted to determine the agreement between participants for all experimental comparative judgements. Results revealed moderate agreement between participant judgements, $\kappa = 0.565$ (95% CI, 0.562 to 0.568), $p < 0.0001$.

Data were then analysed to determine if the comparative judgements were valid: responses to ‘direct instruction’ were used as a measure of face validity as the definition of this feature requires an app to have an avatar that guides and instructs the user (see Table 1) and, hence, can be objectively assessed. If the apps that include an avatar (ie, onebillion, KitKit and CCI) were chosen more frequently as demonstrating ‘direct instruction’ than the apps that did not include an avatar (ie, Chimple and RoboTutor), face validity will be confirmed. Accordingly, a binomial test of probability was conducted using the sum frequency count for ‘direct instruction’ for the three apps that have a clear avatar compared with the two apps that do not use an avatar. The sum frequency of responses to the ‘direct instruction’ feature for the three avatar apps was 333 compared with 77 for the two apps with no avatar. A binomial test revealed that the apps with an avatar were chosen significantly more than chance, $p < 0.001$, demonstrating the comparative judgements made were valid.

Data analyses were then conducted to address each of the research questions posed. For each of the 15 app features examined, participants made 10 comparative judgements, resulting in a maximum frequency count of four per app. As data were not normally distributed, non-parametric tests were conducted. Results are reported at a two-tailed level of probability. Statistical analyses of the results of the experiment were conducted using Jamovi version 1.6 or Python version 3.10.4.

1. Which features characterise each of the five finalist learning apps used in the GLXP?

To identify features that characterised each of the different learning apps, a series of Binomial tests of probability was conducted using sum frequency counts for each feature per app, across the 150 experimental trials. Results indicated app features that were chosen significantly above or below chance.

2. Do the five finalist learning apps of the GLXP, that have been shown to support positive learning outcomes with out-of-school children within the same randomised control trial, share features?

To determine if pedagogical features were common across the five finalist apps a series of 10 Spearman's rank correlation tests were conducted using sum frequencies per app across each of the 15 features examined through the 150 experimental trials. Bonferroni correction was applied to allow for multiple comparisons; adjusted significance level = 0.0125.

RESULTS

1. Which features characterise each of the five finalist learning apps used in the GLXP?

Table 3 reports the sum frequencies of the 15 pedagogical features for each of the five finalist learning apps. Green cells with bold text indicate results from the binomial tests of probability where an app feature was favoured by participants at a level significantly more than chance; red cells with bold text indicate results for app features that were chosen significantly less than chance (at a 5% significance level). For each app, the rank order of features is provided in parentheses.

As shown in Table 3, for the onebillion app (a joint winner of the GLXP) participants judged 13 out of the 15 app features examined as stronger than competitors at a level significantly greater than chance, with only Gamification and Free Play being selected at levels expected by randomness. In contrast, participants preferred all the 15 app features examined for the Robotutor app (which produced the lowest overall learning gains across the GLXP field trial) at a level significantly below chance.

2. Do the five finalist learning apps of the GLXP, that have been shown to support positive learning outcomes with out-of-school children within the same randomised control trial, share features?

As shown in Table 4, Spearman's rho correlation coefficients were calculated to determine the relationship between apps across the 15 pedagogical features examined. Only one correlation was found to be significant when Bonferroni correction was applied for multiple comparisons at an increased significance level of 0.0125. A strong negative correlation was found between two of the runners-up, Chimple and CCI, $r_s(39) = -0.688$, $p = 0.005$, demonstrating these apps differed significantly to one another in pedagogical features.

DISCUSSION

The aim of this study was to identify pedagogical features of interactive apps that are effective in supporting learning of foundational skills for out-of-school children in low-income settings. Here, for the first time, we report findings from a comparative judgement task of the five finalist apps used in the GLXP that directly links app features to learning outcomes established over a 15-month field trial with out-of-school children in Tanzania. Two key findings were revealed.

First, across the five finalist apps, results showed that only the joint winners of the GLXP offered implementations of pedagogical features that were significantly preferred by participants. In contrast, for the three runners-up, participants favoured some pedagogical features (or all features in the case of Robotutor, which showed the least learning gains in the 15-month field trial) significantly less than chance. Accordingly, these results provide evidence that the six pedagogical features shared by onebillion and KitKit—joint winners of the GLXP—are particularly effective in supporting learning of foundational skills with out-of-school children in remote settings—specifically: *autonomous learning, motor skills, task structure, engagement, language demand* and *personalisation*. Three of these features—*autonomous learning, motor skills* and *language demands*—are centred on app accessibility, which has been

TABLE 3 Binomial test of probability results for each of the five finalist learning apps and the 15 pedagogical features.

	Sum frequency count (total across apps = 410; chance = 82) significance of binomial test				
App feature	**onebillion**	**Kitkit**	**CCI**	**Chimple**	**Robotutor**
Direct instruction	**144 (1)** <0.001**	84 (13) 0.844	**105 (3)** 0.007**	**57 (13.5)** 0.002**	**20 (9)** <0.001**
Autonomous learning	**142 (2)** <0.001**	**109 (6)** 0.002**	**62 (13)** 0.013*	82 (4) 1	**15 (13)** <0.001**
Curriculum links	**140 (3)** <0.001**	68 (15) 0.966	98 (6) 0.059	**63 (11)** 0.019*	**41 (1)** <0.001**
Retrieval-based learning	**132 (4)** <0.001**	98 (8) 0.059	92 (7) 0.242	**60 (12)** 0.006**	**28 (7)** <0.001**
Motor skills	**131 (5)** <0.001**	**110 (5)** 0.001**	**46 (15)** <0.001**	90 (3) 0.353	**33 (4)** <0.001**
Task structure	**130 (6)** <0.001**	**124 (2)** <0.001**	85 (9) 0.749	**53 (15)** <0.001**	**18 (11)** <0.001**
Feedback	**130 (7)** <0.001**	92 (9) 0.242	88 (8) 0.493	73 (8) 0.293	**27 (8)** <0.001**
Social interaction	**130 (8)** <0.001**	86 (12) 0.658	**108 (1)** <0.001**	74 (7) 0.355	**12 (14)** <0.001**
Active learning	**129 (9)** <0.001**	83 (14) 0.941	**100 (5)** 0.034*	69 (9) 0.118	**29 (6)** <0.001**
Meaningful learning	**127 (10)** <0.001**	91 (10) 0.294	**102 (4)** 0.019*	**57 (13.5)** 0.002**	**33 (4)** <0.001**
Engagement	**122 (11)** <0.001**	**116 (4)** <0.001**	79 (10) 0.766	77 (6) 0.584	**16 (12)** <0.001**
Language demand	**117 (12)** <0.001**	**107 (7)** 0.003**	72 (11) 0.239	81 (5) 0.961	**33 (4)** <0.001**
Personalisation	**100 (13)** 0.034*	**133 (1)** <0.001**	**107 (2)** 0.003**	**65 (10)** **0.037***	**5 (15)** <0.001**
Gamification	79 (14.5) 0.766	**121 (3)** <0.001**	**59 (14)** 0.004**	**132 (2)** <0.001**	**19 (10)** <0.001**
Free play	79 (14.5) 0.766	88 (11) 0.493	68 (12) 0.091	**140 (1)** <0.001**	**35 (2)** <0.001**

Note: The five finalist apps are ordered according to overall learning gains achieved across the 15-month GLXP field trial, with onebillion and KitKit being joint winners (XPRIZE, 2019). Binomial tests were conducted where n = total number of judgements (trials) made per feature across the study, k = observed sum frequency count for the chosen app per feature, p = 0.2 [the probability that the chosen app will be selected on any particular trial] and q = 0.8 [the probability that the chosen app will not be selected on any particular trial].

$^{**}p<0.01$; $^{*}p<0.05$.

highlighted prominently within the literature (eg, Gulliford et al., 2021; Lynch et al., 2021; Pitchford, 2023). Crompton et al. (2021) highlighted accessibility to be a fundamental barrier to education in low and middle-income countries (LMICs). Our results support prior research indicating that pedagogical app features that increase accessibility are most effective at supporting the acquisition of foundational skills.

Second, results showed no significant positive correlations between the five finalist apps deployed in the GLXP across the 15 features examined in this study, emphasising the variation of features across the applications. Only one correlation was significant, and that was a strong negative correlation, demonstrating a high degree of dissimilarity between the apps

TABLE 4 Spearman's rho correlation matrix between apps across the 15 features examined.

	onebillion	Kitkit	Chimple	CCI	Robotutor
onebillion					
Spearman's rho	—				
p-value	—				
Kitkit					
Spearman's rho	−0.366	—			
p-value	0.180	—			
Chimple					
Spearman's rho	−0.420	0.173	—		
p-value	0.119	0.537	—		
CCI					
Spearman's rho	0.165	−0.418	−0.688**	—	
p-value	0.557	0.123	0.005	—	
Robotutor					
Spearman's rho	0.000	−0.520	0.070	−0.244	—
p-value	1.000	0.047	0.804	0.381	—

**$p < 0.0125$.

produced by Chimple and CCI—each targeting distinctly different sets of features through which to engage users. This variability is somewhat expected, due to the lack of direction given to app developers in the GLXP but is beneficial in the comparisons it allows us to make. The diversity in approaches adopted by the five teams regarding what they thought would constitute an effective learning app in this context is notable, however. For the joint winners of the GLXP, onebillion and KitKit, whose apps resulted in the highest learning gains over the 15-month field trial (XPRIZE, 2019), a weak negative correlation was found across app features, but again this was not significant, reflecting the dissimilarity in pedagogical features within these two apps. This variation between the two winning apps suggests that effective learning of foundational skills through interactive apps can be achieved via a range of pedagogical features but that a core subset of features must be covered well, which relate to *autonomous learning, motor skills, task structure, engagement, language demand* and *personalisation*.

It is not surprising that autonomous learning was shown to be associated with effective learning of foundational skills in this study—as the ability to learn autonomously was almost necessitated—as the children in the GLXP were not provided with formal instruction beyond the app. This corroborates previous research proposing learner autonomy is critical to improving children's motivation, reflective engagement and educational outcomes (Lan, 2018). Apps deployed in out-of-school settings need to prioritise ease of navigation through the content to encourage and motivate independent learning by novice users. This is important even when installing tablets in school settings within LMICs, as there are often low teacher-to-pupil ratios and large class sizes (Jordan et al., 2021). Hence, apps that encompass pedagogical features that promote independent learning are preferable for children to maximise their learning experience.

Results also demonstrated that accessibility in terms of motor skills and language demand is important in fostering positive learning outcomes. Developers need to consider the level of motor skills required for children to interact effectively with their apps: too high level of precision may have a negative impact on accessibility, especially for children with physical disabilities who may not possess the fine motor skills required to interact with the app content (Gulliford et al., 2021; Pitchford et al., 2018). Similarly, the language used in the app needs to be appropriate for the child's developmental age, which could be highly variable in LMIC out-of-school contexts. Language proficiency has been shown to correlate with children's learning with interactive apps (Gulliford et al., 2021; Outhwaite et al., 2020) so developers should keep language as simple as possible to enhance the reach of their apps.

The other three app features shared by the joint winners of the GLXP were task structure, engagement and personalisation. The importance of task structure has been highlighted in previous research with app use in primary school settings, due to the complementary relationship it has with the curriculum and the use of reinforcement when learning new topics (Gulliford et al., 2021). Likewise, personalisation (with programmatic levelling), in combination with feedback, has been suggested to maximise learning outcomes when considering literacy app design (Vanbecelaere et al., 2020) and maths app interventions (Outhwaite et al., 2022). For an app to have high educational quality, it should also support a child's engagement in the learning process, as engagement and learning 'go hand in hand' (Raymer, 2013), using contingent interactions, extrinsic and intrinsic motivation (Hirsh-Pasek et al., 2015).

Interestingly, Free Play as a feature did not significantly characterise the two winning apps of the GLXP, despite previous research suggesting it plays an important role in exploration and autonomy (Hirsh-Pasek et al., 2015). Play-based learning (or guided play) has recently been shown to be successful in school settings (within a developed context), suggesting that play can be valuable if guided by a teacher with a learning objective, by balancing exploration and instruction (Skene et al., 2022). The importance of task structure highlighted in this study corroborates Skene et al. (2022) and demonstrates that in *out-of-school* settings, a certain level of structure is required within an app to direct a child through its content and to promote positive learning outcomes where there is no formal teacher available to guide the learning process.

Limitations

A potential limitation of this study is that participants were based in the United Kingdom, with no experience of or affiliation with Tanzania, where the GLXP was undertaken. Participants were thus not familiar with the context in which the apps were deployed which might be particularly pertinent when judging features such as 'meaningful learning' and 'curriculum links' for children in Tanzania. However, the apps investigated in this study are used in other countries and contexts, such as the United Kingdom and Brazil, in which demonstrable learning gains have been achieved (Outhwaite et al., 2020), indicating their effectiveness is not country specific.

CONCLUSION

This study has identified six pedagogical app features—*autonomous learning, motor skills, task structure, engagement, language demand* and *personalisation*—that are significantly associated with learning foundational skills in low-income community settings. This combination of app features appears to be key to ensuring the optimal effectiveness of learning apps deployed in LMICs, where spending budgets for education are extremely limited. Future studies should assess the reliability of this combination of pedagogical features in other educational apps and settings to evaluate their adoption in different contexts.

Effectiveness of app-based learning in LMIC remote community-based settings is also likely to be influenced by how educational apps are implemented, as implementation is critical in determining learning gains with educational apps in school-based settings (eg, Outhwaite, Gulliford, & Pitchford, 2019). Factors pertaining to the environment in which the child resides, as well as factors attributable to the child, and to government education policy, may also influence learning outcomes with educational apps, as discussed by Pitchford (2023). For a comprehensive understanding of how children learn foundational skills with educational apps, research needs to investigate each of these potential influences and then synthesise findings across studies.

The results of this study should inform pedagogical design of educational apps, particularly for use by children of primary school age in LMICs and should be useful to governments, educators and parents, when deciding on educational apps to support the acquisition of foundational skills, especially with out-of-school children. This is crucial considering that 244 million children worldwide are estimated to be out-of-school (UNESCO, 2022) and 10% of global spending is purported to be wasted on 'poor education' that perpetuates significant inequalities in access to and provision of quality education that is failing to produce the desired learning outcomes for children (UNESCO, 2019). Good quality educational apps, with the combination of pedagogical features identified in this study, could start to resolve this global crisis in foundational learning.

ACKNOWLEDGEMENTS

We wish to thank the participants who gave up their time to take part in this study. This work was supported by the XPRIZE Foundation and the Economic and Social Research Council (ESRC) [grant number ES/J500100/1]. We would like to thank the XPRIZE Foundation for supporting this research through donations of the app hardware, the ESRC for funding the research and the finalist teams for ensuring continued access to the software.

CONFLICT OF INTEREST STATEMENT

This work was undertaken as part of a doctoral research programme in collaboration with the XPRIZE Foundation. The XPRIZE Foundation facilitated access to the apps and tablets that were used in the competition but otherwise did not influence the research design, data acquisition process, data analysis or interpretation of research findings. Accordingly, 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.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Ethical approval for this study was granted by the University of Nottingham Ethics Committee [ref: S1377].

ORCID

Bethany Huntington https://orcid.org/0000-0003-0437-2090
James Goulding https://orcid.org/0000-0002-8892-6398
Nicola J. Pitchford https://orcid.org/0000-0001-5193-8371

REFERENCES

Al-Azawi, R., Al-Faliti, F., & Al-Blushi, M. (2016). Educational gamification vs. game-based learning: Comparative study. *International Journal of Innovation, Management and Technology*, *7*(4), 132–136.

Alemán de la Garza, L., Anichini, A., Antal, P., Beaune, A., Crompton, H., & Tsinakos, A. (2019). *Rethinking pedagogy. Exploring the potential of digital technology in achieving quality education*. UNESCO MGIEP.

Allison, K. (2019). *An evaluation of tablet-based apps for maths learning in young children with Down's syndrome* [Unpublished doctoral dissertation]. University of Nottingham.

Azevedo, J. P., Hasan, A., Goldemberg, D., Geven, K., & Iqbal, S. A. (2021). Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. *The World Bank Research Observer*, *36*(1), 1–40. https://doi.org/10.1093/wbro/lkab003

Benton, L., Mavrikis, M., Vasalou, A., Joye, N., Sumner, E., Herbert, E., Revesz, A., Symvonis, A., & Raftopoulou, C. (2021). Designing for "challenge" in a large-scale adaptive literacy game for primary school children. *British Journal of Educational Technology*, *52*(5), 1862–1880. https://doi.org/10.1111/bjet.13146

Bettinger, E., Fairlie, R. W., Kapuza, A., Kardanova, E., Loyalka, P. K., & Zakharov, A. (2020). *Does EdTech substitute for traditional learning? Experimental estimates of the educational production function* (Working Paper No. 26967). National Bureau of Economic Research. https://doi.org/10.3386/w26967

Bramley, T. (2007). Paired comparison methods. In P. E. Newton, J. Baird, H. Goldstein, H. Patrick, & P. Tymms (Eds.), *Techniques for monitoring the comparability of examination standards* (pp. 246–294). Qualifications and Curriculum Authority.

Callaghan, M. N., & Reich, S. M. (2018). Are educational preschool apps designed to teach? An analysis of the app market. *Learning, Media and Technology*, *43*(3), 280–293. https://doi.org/10.1080/17439884.2018.1498355

Chodura, S., Kuhn, J. T., & Holling, H. (2015). Interventions for children with mathematical difficulties: A meta-analysis. *Zeitschrift fur Psychologie*, *223*, 129–144. https://doi.org/10.1027/2151-2604/a000211

Crompton, H., Chigona, A., Jordan, K., & Myers, C. (2021). *Inequalities in girls' learning opportunities via EdTech: Addressing the challenge of Covid-19* (Working Paper 31). EdTech Hub. https://doi.org/10.5281/ZENODO.4917252

Fisher, D., & Frey, N. (2021). *Better learning through structured teaching: A framework for the gradual release of responsibility*. ASCD.

Forbes. (2021, August 19). *Education apps kept kids learning for over a year, now these apps will shape our future*. https://www.forbes.com/sites/eladnatanson/2021/08/19/education-apps-kept-kids-learning-for-over-a-year--now-these-apps-will-shape-our-future/?sh=3690724f4d35

Grimaldi, P. J., & Karpicke, J. D. (2014). Guided retrieval practice of educational materials using automated scoring. *Journal of Educational Psychology*, *106*(1), 58–68. https://doi.org/10.1037/a0033208

Gulliford, A., Walton, J., Allison, K., & Pitchford, N. (2021). A qualitative investigation of implementation of app-based Maths instruction for young learners. *Educational & Child Psychology*, *38*(3), 90–108.

Herodotou, C. (2018). Young children and tablets: A systematic review of effects on learning and development. *Journal of Computer Assisted Learning*, *34*(1), 1–9. https://doi.org/10.1111/jcal.12220

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3–34. https://doi.org/10.1177/1529100615569721

Huntington, B., Goulding, J., & Pitchford, N. (2021). Transforming global learning with digital technologies: A qualitative exploration of the use of educational technology with marginalised, out-of-school children living in remote settings. In *EDULEARN21 proceedings* (p. 9201). IATED. https://doi.org/10.21125/edulearn.2021.1853

Jordan, K., Proctor, J., Koomar, S., & Bapna, A. (2021). *A country-level research review: EdTech in Tanzania* (Working Paper). EdTech Hub. https://doi.org/10.5281/zenodo.4618376

Kalton, G., & Schuman, H. (1982). The effect of the question on survey responses: A review. *Journal of the Royal Statistical Society Series A (General)*, *145*, 42–73. https://doi.org/10.2307/2981421

Kim, J., Gilbert, J., Yu, Q., & Gale, C. (2021). Measures matter: A meta-analysis of the effects of educational apps on preschool to grade 3 children's literacy and math skills. *AERA Open*, *7*(1), 1–19. https://doi.org/10.1177/23328584211004183

Kolak, J., Norgate, S. H., Monaghan, P., & Taylor, G. (2021). Developing evaluation tools for assessing the educational potential of apps for preschool children in the UK. *Journal of Children and Media*, *15*(3), 410–430. https://doi.org/10.1080/17482798.2020.1844776

Lan, Y. J. (2018). Technology enhanced learner ownership and learner autonomy through creation. *Educational Technology Research and Development*, *66*(4), 859–862. https://doi.org/10.1007/s11423-018-9608-8

Lee, C. Y., & Cherner, T. S. (2015). A comprehensive evaluation rubric for assessing instructional apps. *Journal of Information Technology Education*, *14*, 21–53. http://www.jite.org/documents/Vol14/JITEV14ResearchP021-053Yuan0700.pdf

Lee, H. M., & Loo, P. A. (2021). Gamification of learning in Early age education. *Journal La Edusci*, *2*(2), 44–50. https://doi.org/10.37899/journallaedusci.v2i2.380

Lynch, P., Singal, N., & Francis, G. (2021). *EdTech for learners with disabilities in primary school settings in LMICs: A systematic literature review* (Working Paper). EdTech Hub. https://doi.org/10.5281/zenodo.4348995

Major, L., & Francis, G. (2020). *Technology-supported personalised learning: Rapid evidence review*. EdTech Hub. https://edtechhub.org/wp-content/uploads/2020/09/Rapid-Evidence-Review_-Technology-supported-personalised-learning.pdf

Marshall, N., Shaw, K., Hunter, J., & Jones, I. (2020). Assessment by comparative judgement: An application to secondary statistics and English in New Zealand. *New Zealand Journal of Educational Studies*, *55*(1), 49–71. https://doi.org/10.1007/s40841-020-00163-3

Northrop, L., & Killeen, E. (2013). A framework for using iPads to build early literacy skills. *The Reading Teacher*, *66*(7), 531–537. https://doi.org/10.1002/trtr.1155

Outhwaite, L., Early, E., Herodotou, C., & Van Herwegen, J. (2022). *Can Maths apps add value to young children's learning? A systematic review and content analysis*. Nuffield Foundation.

Outhwaite, L. A., Faulder, M., Gulliford, A., & Pitchford, N. J. (2019). Raising early achievement in math with interactive apps: A randomized control trial. *Journal of Educational Psychology*, *111*(2), 284–298. https://doi.org/10.1037/edu0000286

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 and Method in Education*, *43*(3), 225–242. https://doi.org/10.1080/1743727X.2019.1657081

Outhwaite, L. A., Gulliford, A., & Pitchford, N. J. (2020). Language counts when learning mathematics with interactive apps. *British Journal of Educational Technology*, *51*(6), 2326–2339. https://doi.org/10.1111/bjet.12912

Papadakis, S., & Kalogiannakis, M. (2017). Mobile educational applications for children: What educators and parents need to know. *International Journal of Mobile Learning and Organisation*, *11*(3), 256–277. https://doi.org/10.1504/IJMLO.2017.085338

Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2017). Designing and creating an educational app rubric for preschool teachers. *Education and Information Technologies*, *22*(6), 3147–3165. https://doi.org/10.1007/s10639-017-9579-0

Papadakis, S., Kalogiannakis, M., & Zaranis, N. (2018). Educational apps from the android Google play for Greek preschoolers: A systematic review. *Computers & Education*, *116*, 139–160. https://doi.org/10.1016/j.compedu.2017.09.007

Pitchford, N. J. (2015). Development of early mathematical skills with a tablet intervention: A randomized control trial in Malawi. *Frontiers in Psychology*, *6*(485), 1–12. https://doi.org/10.3389/fpsyg.2015.00485

Pitchford, N. J. (2023). Customised E-learning platforms. In T. Madon & A. J. Gadgil (Eds.), *Introduction to development engineering: A framework with applications from the field* (pp. 269–292). Springer. https://doi.org/10.1007/978-3-030-86065-3

Pitchford, N. J., Kamchedzera, E., Hubber, P. J., & Chigeda, A. (2018). Interactive apps promote learning in children with special educational needs. *Frontiers in Psychology*, *6*, 262–274. https://doi.org/10.3389/fpsyg.2015.00485

Pitchford, N. J., Papini, C., Outhwaite, L. A., & Gulliford, A. (2016). Fine motor skills predict maths ability better than they predict reading ability in the early primary school years. *Frontiers in Psychology*, *7*, 783. https://doi.org/10.3389/fpsyg.2016.00783

Plump, C. M., & LaRosa, J. (2017). Using Kahoot! in the classroom to create engagement and active learning: A game-based technology solution for eLearning novices. *Management Teaching Review*, *2*(2), 151–158. https://doi.org/10.1177/2379298116689783

Pollitt, A. (2012). The method of adaptive comparative judgement. *Assessment in Education: Principles, Policy & Practice*, *19*(3), 281–300. https://doi.org/10.1080/0969594X.2012.665354

Putra, A. S., Warnars, H. L. H. S., Abbas, B. S., Trisetyarso, A., Suparta, W., & Kang, C. H. (2018, September). Gamification in the e-learning process for children with attention deficit hyperactivity disorder (ADHD). In *2018 Indonesian Association for Pattern Recognition International Conference (INAPR)* (pp. 182–185). IEEE.

Raymer, R. (2013, September 13). The rock stars of eLearning: An interview with Karl Kapp. *eLearn Magazine*. http://elearnmag.acm.org/archive.cfm?aid=2524223

Richards, J. C. (2015). Technology in language teaching today. *Indonesian JELT*, *10*(1), 18–32. https://doi.org/10.25170/ijelt.v10i1.1506

Skene, K., O'Farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., & Ramchandani, P. G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development*, *93*(4), 1162–1180. https://doi.org/10.1111/cdev.13730

Statista. (2021, July 6). *Mobile education app downloads worldwide*. https://www.statista.com/statistics/1128262/mobile-education-app-downloads-worldwide-platforms-millions/

Stubbé, H., Badri, A., Telford, R., van der Hulst, A., & van Joolingen, W. (2016). E-learning Sudan, formal learning for out-of-school children. *Electronic Journal of e-Learning*, *14*(2), 136–149.

Ting, Y. L. (2015). Tapping into students' digital literacy and designing negotiated learning to promote learner autonomy. *The Internet and Higher Education*, *26*, 25–32. https://doi.org/10.1016/j.iheduc.2015.04.004

Toub, T. S., Rajan, V., Golinkoff, R. M., & Hirsh-Pasek, K. (2016). Guided play: A solution to the play versus learning dichotomy. In D. C. Geary & D. B. Berch (Eds.), *Evolutionary perspectives on child development and education* (pp. 117–141). Springer. https://doi.org/10.1007/978-3-319-29986-0

UNESCO. (2019). *GEM Report: Migration, displacement and education—Building bridges, not walls*. https://gem-report-2019.unesco.org/

UNESCO. (2022, September). *New estimation confirms out-of-school population is growing in Sub-Saharan Africa* [Policy Paper]. https://www.unesco.org/gem-report/en/2022-out-school#:~:text=It%20is%20estimated%20that%20244,and%20125.5%20million%20were%20boys.

Vanbecelaere, S., Cornillie, F., Sasanguie, D., Reynvoet, B., & Depaepe, F. (2021). The effectiveness of an adaptive digital educational game for the training of early numerical abilities in terms of cognitive, noncognitive and efficiency outcomes. *British Journal of Educational Technology*, *52*(1), 112–124. https://doi.org/10.1111/bjet.12957

Vanbecelaere, S., Van den Berghe, K., Cornillie, F., Sasanguie, D., Reynvoet, B., & Depaepe, F. (2020). The effectiveness of adaptive versus non-adaptive learning with digital educational games. *Journal of Computer Assisted Learning*, *36*(4), 502–513. https://doi.org/10.1111/jcal.12416

XPRIZE. (2019). *Global learning XPRIZE executive summary*. https://www.xprize.org/prizes/global-learning/articles/glexp-executive-summary

How to cite this article: Huntington, B., Goulding, J., & Pitchford, N. J. (2023). Pedagogical features of interactive apps for effective learning of foundational skills. *British Journal of Educational Technology*, *00*, 1–19. https://doi.org/10.1111/bjet.13317