Evaluating performance on a bespoke maths game with children with Down syndrome

Received: 3 August 2021 | Revised: 28 February 2022 | Accepted: 26 April 2022
DOI: 10.1111/jcal.12685

ARTICLE

Journal of Computer Assisted Learning WILEY

Evaluating performance on a bespoke maths game with children with Down syndrome

Jill Porter

Institute of Education, University of Reading, Reading, UK

Correspondence
Jill Porter, Institute of Education, University of Reading, Reading, UK.
Email: j.porter@reading.ac.uk

Abstract

Background: Despite the interest and potential of multi touch devices, there are limited published studies researching their effectiveness and usability specifically with children with Down syndrome, one of the most common groups of children with an intellectual disability. This is particularly true for mathematical learning, an area in which many experience particular difficulty.

Objectives: The study set out to evaluate a bespoke digital game in which children learned to select which was "more," a foundational skill for understanding magnitude.

Methods: A mixed methods approach was adopted with eight single case studies of children aged 9–14 years. Probes (untaught examples) were introduced to examine progress and the impact on performance in non-digital contexts.

Results: Five pupils improved their performance on the digital games and this was sustained at the time of the delayed post-test. Four pupils showed improved performance in non-taught, non-digital contexts for both taught and untaught ratios. Disaggregated data reveals the variability in performance, with peak performances occurring at different points of the intervention. The introduction of a two-player version improved performance for five pupils through promoting sustained attention and strategic responses to winning. One child performed at chance level on the digital game throughout but made gains in non-digital settings.

Conclusions: Bespoke learning tools have good potential to promote attention to numerosity. However, progress in digital contexts does not automatically transfer to non-digital contexts. The study reveals the individual nature of the learning affordances of different pedagogic tools and the place of bespoke games within teachers' repertoires.

KEYWORDS
bespoke serious game, Down syndrome, magnitude, numerosity, tablet

1 | INTRODUCTION

There is a growing interest in the use of digital games for children with special needs, including those with an intellectual disability (Tsikinas & Xinogalos, 2019). The use of touch screen has made these particularly accessible, but with the quality of the software being key in enabling educational goals to be met as well as promoting independent learning (Feng et al., 2010). Evidence of the effectiveness of digital games in bringing about new learning and usability is mixed, in part a reflection of the methodological challenges of research with atypical populations (Nacher et al., 2018). There have also been changes towards higher expectation of mathematical achievements of pupils with intellectual disabilities over the last 10 years (Spooner et al., 2019) and recognition that interventions need to be tailored to the strengths and needs of particular groups (Lemons et al., 2015: King et al., 2017). The use of multimedia materials with children with Down syndrome (DS) are argued to be particularly appropriate learning approach. Children with DS can experience a number of challenges that impact on their learning albeit to differing degrees. They are more likely than their peers to experience visual difficulties with poor near vision (Cregg et al., 2001) and have an auditory impairment (Park et al., 2012). These in turn impact on the processing of information. Visual processing is generally stronger than auditory processing although this does not hold true for every context (Baddeley & Jarrold, 2007; Yang et al., 2014). Children are more likely to experience difficulties with fine motor movements (Fidler & Rogers, 2006) all of which impact on their interaction with technology. Researchers have also drawn attention to qualitative differences that they bring to learning tasks including variable performance, and avoidant strategies (Wishart, 2005) being distractible and having difficulties in attention and motivation (Fidler & Rogers, 2006). These difficulties are evidenced in difficulty learning to count where pupils struggle to acquire and produce a stably ordered list of number words and apply them whilst tagging objects to count objects (Onnivello et al., 2019). This has led a number of researchers to suggest that too much emphasis is placed on counting and alternative approaches should be sought (Clarke & Faragher, 2014). Digital games can offset some of these challenges to learning. They can be visually attractive, allow for individualisation and sequencing of tasks, provide repetition to help retention, promoting attention and motivation (Felix et al., 2016; Ortega-Tudela & Gómez-Arizaw, 2006). Their potential has been explored in a few small scale studies that have targeted counting and numeracy recognition (Agheana & Duta, 2015; Ahmad et al., 2014; Ortega-Tudela & Gómez-Arizaw, 2006) with a limited number designed specifically for this pupil group (Shafie et al., 2013) and mostly favouring an online work card approach.

The lack of bespoke serious games for children with DS reflects the complexity of bringing together different areas of expertise. Their design demands close collaboration between a number of different professionals: digital game and software designers, professionals with a pedagogic understanding of the focal curriculum area and a knowledge of the ways in which the strengths and needs of different special educational needs impact on learning. It is therefore not surprising that the majority of studies use existing mainstream games with children with intellectual disabilities. The study reported here is of a bespoke digital game designed and developed through two cycles of activity in collaboration with colleagues from a Department of Computer Science, to promote an awareness of magnitude (or numerosity) in children with DS. The focus of this paper is the second cycle, the first concerned the technical performance of the App with a single performance of nine young people. This second cycle explored in more depth the utility of the App to promote learning and how that might best be evaluated. The aim was to develop an ethical methodology for evaluation through which to test out the effectiveness of the digital in bringing about new learning in the understanding of magnitude.

2 | THE TARGET SKILL

Here we focus on the ability to discriminate between quantities, of knowing which is more, skills that are more typically developed in the pre-school period. Representation of magnitude has been argued to lie at the core of mathematical development (Siegler & Lortie-Forgues, 2014) supporting the development of precise representations and the developmental shift from non-symbolic to symbolic representations of number. Developing a game that strengthens the ability to focus on numerosity has the potential to provide a strong platform on which to build a range of mathematical skills. An important aspect of the design is the way we introduce challenge. In non-digital contexts, the measure of progression is usually taken as the ratio between two quantities for larger sets that are outside the subitizing range, so that quantities that differ in the ratio 1:3 (e.g., 3:9; 4:12) are easier and acquired much earlier than those of 3:4 (e.g., 6:8; 8:12). Research indicates that children with DS follow a developmental progression in being able to discriminate between quantities with evidence to suggest that, in least in some contexts, children may find tasks involving small set sizes more difficult than those with larger set sizes (Porter, 2019; Onnivello et al., 2019). The challenges introduced in the game need to reflect these sequences.

Traditionally, in non-digital settings, children's skills are tested by repeated presentation to two arrays, often but not always, dots, requiring a point or touch response to the one which has more. These studies of children with DS's abilities often required sustained attention, with sessions lasting as much as 50 minutes (Abreu-Mendoza & Arias-Trejo, 2015) and frequently deploy large numbers of presentations (as many as 103 in Sella et al., 2013). These "test" situations can be contrasted to the development of (serious) games where the focus is on maximizing children's engagement and motivation to succeed. We know from non-digital studies that presentation is important. Linear board games, as opposed to those which involve a circular track, result in better scores on magnitude comparisons, estimating and addition (Siegler & Ramani, 2009). Comparisons which involve simultaneous presentations can produce different results to those which are sequential (Yang et al., 2014) and there is some evidence that children with DS do better when information is presented in a particular part of the screen rather than being dispersed across it (Belacchi et al., 2014). Further design guidance can be found through looking specifically at the studies that evaluate children's acquisition of magnitude representation and comparison using digital games.

3 | MAGNITUDE REPRESENTATION AND DIGITAL GAMES

To date, two game based studies with children with DS have focused on magnitude. Porter (2016, 2018) reports on a study of a collaboratively designed game for an ipad described as a supplement rather than a replacement for more traditional activities. Awareness of magnitude was used functionally and the game mechanic was developed working with a group of young mainstream coaches, and evaluated with a group of children with DS (Porter, 2018). This mixed methods study captured a snapshot of over 64 children's responses on being introduced to the game over three iterations of its development. It illustrates the variability in performance within this group of children and the requirement for a multifaceted approach to understanding progression and levels of difficulty. Length of game, level of challenge and pace of change, requirement for sustained attention all contributed to the relative difficulty of a game.

A series of issues were highlighted by Porter, 2018's evaluation: a number of children responded better in the non- digital equivalent of the game, while the children were clearly motivated to interact with the game, "engagement proved a mixed blessing" (p436) with children keen to explore elements that distracted them from the goal of the game. This led to further game development; some extension to the levels, making an easier entry level requiring shorter play time, and additional digital rewards for completion of the game. Together these helped to counteract children disengaging when they encountered failure, but the impact on learning was not fully tested.

More recently Sella et al. (2021) report on an experimental study using the mainstream game "The Number Race" with 30 children with DS, and a similar size control group who worked with software designed to improve reading skills. The experimental group showed improvements in several but not all number tasks at the end of 20 weeks of training. Sella et al. (2021) conclude on the advantages of targeting a single skill rather than the many which were addressed in the number race. The authors illustrate how the number race programme has a differential impact on two children with DS depending on individual strengths and difficulties and their starting points (Onnivello et al., 2019). As they report "Sometimes, as a joke, he would deliberately make mistakes or wait until the last moment to answer, especially when he was becoming bored," (Onnivello et al., 2019, p. 281). This reinforces the importance of knowing when to stop, attention to motivation, and the fundamental contribution of qualitative data to an evaluation study. Programmes have to accommodate to a range of abilities and attitudes towards the task.

Thus, while the *potential* of serious digital games with this group are well recognized (Ortega-Tudela & Gómez-Arizaw, 2006) the findings are mixed. There is however a slowly growing body of data that indicates the importance of attending to a number of key features and adjusting the pedagogy to the characteristics of the group. Nascimento et al. (2020) identifies the importance of creating a strong narrative together with a compelling goal. This can be a missing feature when online materials mimic those of traditional lessons, and in the case of children and young people with an ID it often results in the use of kindergarten materials. These do not provide the developmental pace that is appropriate for many children with DS (Feng et al., 2010).

Accessibility is facilitated through the use of touch screens which are more intuitive than other interactive devices but these usually require a range of different touch movements. The specifies of the interface are an important design consideration. A few researchers have explored the relative difficulty for people with DS. Nacher et al. (2018) describe the frustration and disenchantment when imprecise movements had unintended consequences for their primary aged children with DS. Martin-Gutierrez and Del Rio Guerra (2021) provide a comparative list for different gestures- with higher success rates for tap, touch and hold, stretch, slide and separate, drawing on an analysis of gesture completion times, success rate and object size on a purpose-built gaming App. Their qualitative data also indicates the importance of getting this interface right with participants getting very frustrated when they stopped the gesture too soon and were unable to finish the task; and when a double tap response led to errors.

In summary the literature reveals the importance of careful design and development of digital games if they are not to have unintended consequences. This includes:

- An analysis of the target skill, levels of difficulty or progression, and pace of learning;
- The specificities of the interface, visual array and response type;
- The accompanying narrative and support structures;
- Careful attention to what will motivate the child to engage and perform at their best;
- Session lengths that do not exceed the interest of the child.

4 | METHODOLOGICAL ISSUES IN ESTABLISHING VALIDITY IN EVALUATION STUDIES

The evaluation literature for digitally based interventions with children with DS, or even the broader based group of children with ID, is relatively sparse with much reported briefly as conference papers where methodology and its challenges are given limited space. A review by Tsikinas and Xinogalos (2019) indicates that many studies are small scale and use a pre-test post-test design. Many like Shafie et al. (2013) could be viewed (as indeed is this study) as developmental work rather than adhering closely to experimental single case study designs of the kind described in detail by Horner et al. (2005); and subsequently formally set out by the Council of Exceptional Children (Ledford et al., 2018). Targeting a particular SEN group fosters a recognition of the ways in which particular syndromic characteristics impact on both the focal topic but also on more common learning characteristics, although in the field of mathematics this is less common than in other curricula areas (King et al., 2017; Lemons et al., 2015). However, it can also perpetuate assumptions of group level commonalities at the expense of "personal patterns of learning" (Felix et al., 2016). Aggregated group data often camouflages the individual variation that is evident in every area of development, (Karmiloff-Smith et al., 2016). Comparing groups, those who receive the intervention and those who receive an alternative requires the establishment of equivalence through some kind of matching procedure. Given the degree of individual differences, this requires access to a large pool of potential participants and strict inclusion criteria. Many studies use convenience samples. It is therefore logical to have each participant serving as their own control and, as Horner et al. (2005) suggest, identify performance patterns before and after intervention. The advantage of digital interventions in this context is that they provide a standardized intervention, thereby addressing issues of fidelity and provide a reliable recording system for measuring responses.

Experimental case study designs require a stable baseline. This raises a number of dilemmas for the design of the methodology. A number of writers have commented on the instability of responses from DS children (Hasan & Messer, 1997; Wishart, 2005), Researchers often use staggered or multiple baselines, as children join the intervention in turn, as a way of controlling for the effects of activities external to the intervention that might be bringing about changes in the behaviour. This strategy has merit when children take part from a single class but is a less convincing method for addressing issues of internal validity when the participants come from different classes. Further, the frequent presentation of "to-be-learned" material raises ethical as well as methodological issues given the affective responses reported in the data (Porter, 2018). Methods that expose children to repeated exposure to failure contribute to promoting disengagement in the face of challenging circumstances, reducing self-confidence and the motivation to succeed.

The use of probes, markers of whether there has been change in the target behaviour, offers an alternative where for example extended multiple base-lines prove impractical (Horner & Baer, 1978). The introduction of probes, untaught examples, can help to strengthen the assessment of the relationship between the intervention and outcome and was adopted here.

To further strengthen the validity, the efficacy of a digital programme can be judged by comparing the learning outcomes with progress on other activities. In the area of mathematics, context can shape pupils understanding and their responses to a task, abstract concepts can acquire new meanings and functions as well illustrated by the teaching of algebra (Monari Martinez & Pellegrini, 2010). In particular the meaning of the task may change. This may help or hinder performance. Social context, the presence of a more able partner, can provide a model but also a motivator as illustrated by Porter (2018) in developmental work.

In addition to efficacy are issues of utility or the usefulness of the learning. If learning is specific to that context it has limited utility. This is a particular issue with respect to mathematics and young people with an intellectual difficulty where there is a sharpened focus on functional skills. The purpose here is not simply to teach children to be successful on the game, but to heighten awareness and attention to differences in numerosity. Transfer of these skills includes being able to differentiate between numerosities of the taught ratios in new contexts and possibly to untaught ratios. Including probes that are untrained but require related skills as part of the design provides the potential for revealing generalization (Browder et al., 2012).

Mindful of these issues of validity and ethics, it was therefore decided to use single case studies, introducing probes during the course of the intervention and to examine progress pre and post (with a follow-up), and the impact on performance in non-digital contexts. It was important that children (and not just their parents) actively opted in to the research, that information was accessible, and that consent was seen as an ongoing process (Porter, 2014). Careful consideration was given to the frequency and duration of the intervention procedure, balancing the importance of consolidating learning with issues of boredom. Lemons et al. (2015) in a review of maths intervention with this group identified large variation both in session length (15–180 minutes) length of intervention (1 session to 6 months) and frequency (1–4 times a week). Sella et al. (2021) argued in the specific context of digital games for the need for more than more than 20 twice weekly sessions to find a substantial difference between groups. The parameters for these decisions will also be determined by pragmatic issues of access.

5 | THE STUDY

The research aimed to examine children with Down syndromes responses to a bespoke game. More specifically it was guided by the following research questions:

> **RQ1.** *Do children's performance on identifying "more" improve through playing digital games? This was operationalized as:*
>
> **RQ1a.** *Do they improve on the baseline pre-test performance measure at the end of the intervention period and;*
>
> **RQ1b.** *Is this sustained at follow-up?*
>
> **RQ2.** *Does their learning generalize to new contexts? This was operationalized as:*
>
> **RQ2a.** *Do the gains transfer to new non-digital contexts?*
>
> **RQ2b.** *Do they improve on non-taught (harder) ratios?*

Following ethical approval from the University, parents of 10 children with Down syndrome from one special school were approached to take part in the study. Parents of eight children returned their consent forms. Those eight children were then invited to take part and given easy read, word and symbol, explanations of the project. In addition, for each session, children's willingness to continue was verbally checked and the activities for the session and expectations clearly set out. Using tangible markers of putting robots in pots, children were able to monitor how many games they had done and how many more before the session ended. Additionally, children had their own booklet for putting in stickers to denote sessions completed. One child chose not to join the session on one occasion, preferring to stay in class to participate in that activity.

5.1 | Participants

Eight children (4 boys and 4 girls) with DS, ranging in chronological age between 9 years 6 months and 14 years 6 months with a mean age of 12 years 3 months participated in this study, working at a small desk with two researchers outside their classroom. Their visual acuity was screened using the Kay Pictures Test (crowded logMar Kay Picture) prior to being assessed in areas related to awareness of more: counting and its components, ability to subitize using a conventional array, and order numbers.

As Table 1 reveals the children showed different levels of ability on the tested skills; the ability to recite numbers in correct order varied between 4 and 22; the ability to count objects varied between 3 and 10; and when requested to give a number of items, this varied between 2 and 5.

When tested on their ability to subitize using cards with arrays of dots in conventional dice patterns only two were able to label all six patterns. Their knowledge of more was also tested through requests to order cards with different dot patterns. They were presented with cards of 1, 2 and 3 dots; 3, 4 and 5 dots; 1, 4 and 8 dots; and 3, 8 and 12 dots. Responses to this task varied from 0 to 4 correct.

5.2 | Design of the study

Intervention was carried out over a 4 week period with three sessions a week lasting approximately 20 min and a follow-up 2 weeks later (see Table 2). Each child formed their own control with probes of progress after three sessions during which they played the game nine times, and again after a further four sessions. The probe conditions involved non-digital activities of a card game and equivalent board game. After four sessions, children were given access to variations of the digital game, with children choosing the version they wanted to play. These included Millie Moreorless, an alternative digital magnitude game and a two player version of Anna Robot. One child who continued to interact at the level of chance was introduced in this second half to non-digital games. The children completed between 212 and 352 (mean 280) discriminations of "more" on digital games.

5.3 | Activities

5.3.1 | Anna's robot

The game was developed for Android tablets using Android Studio. The narrative is the child is helping a scientist who is going to his laboratory to build a robot. In order to ensure an appropriate level of challenge corresponding to each player's ability, the game records and responds to the choices made by adjusting its difficulty. The starting ratio for first and second level are 1:3 and 1:2 respectively. As the player consistently makes correct choices, the game increases its difficulty level by switching to the harder ratio. Conversely, if the player struggles with a particular ratio, the difficulty is decreased Children were invited to play a tablet game comprising of three windows. In the first, (see Figure 1) two square boxes are presented as a dice equivalent but with random presentation of dots in the ratios of 1:3 (1 and 3; 3 and 9 or 4 and 12).When the correct one is chosen it lights up, and the player is required to move each dot in turn to a path to form stepping stones for the scientist towards the lab. An incorrect selection results in a cross appearing

TABLE 2 Design of the study

	Activity
Baseline	Card-game
Time 1–3	Digital Game – Anna Robot
Transfer Probe 1	Nondigital equivalent board game and card game
Time 4	Digital Game – Anna Robot
Time 5, 6	Digital game – Anna Robot 2 and Millie Moreorless
Time 7	Post- test Digital Game – Anna Robot
Transfer Probe 2	Nondigital equivalent board game and card-game
Follow-up	Baseline-Cards and Digital Game-Anne Robot

TABLE 1 Children's responses to number sense tasks

Name	Age	Recite number string (max = 30)	Count objects (max = 10)	Give objects (max = 5)	Label dice pattern (max = 6)	Order cards (max = 4)
Ginny	13 years 8 months	20	10	3	6	3
Fred	12 years 7 months	13	5	5	6	4
Jo	12 years 8 months	22	10	5	3	3
Rizzi	12 years	8	5	5	4	3
Frank	14 years 6 months	10	10	2	3	0
Kat	11 years	4	3	2	4	1
Adi	9 years 6 months	11	10	4	2	0
Jonnie	11 years 9 months	15	10	4	3	2

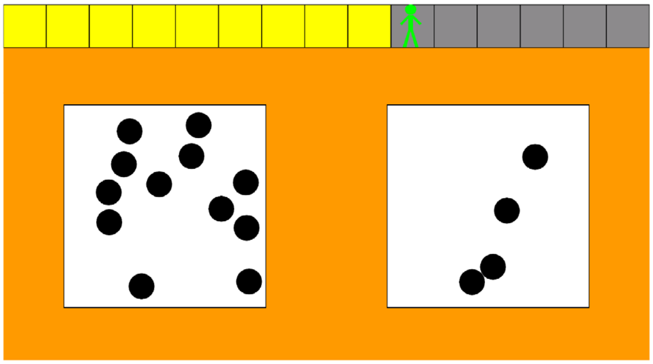

FIGURE 1 The scientist goes to the lab

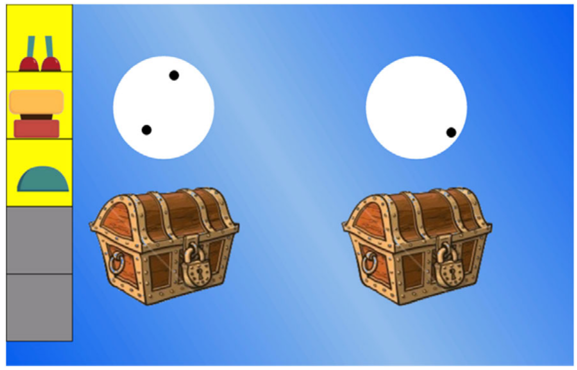

FIGURE 2 Finding the robot parts

to denote that it is wrong (the dots cannot be moved and the scientist cannot progress along the path) and two new quantities are presented. After a number of correct choices the scientist reaches the end of the path to the lab, a second window opens (Figure 2) where the task is to choose between two circular arrays of random dots in the ratios of 1:2 (1 and 2; 2 and 4, and 4 and 8), a correct choice unlocks a chest and part of a robot is presented. The player drags the robot part onto a shelf on the side of the screen and when all five parts have been collected, the final screen opens with all five robot parts for the child to assemble by dragging them together, at which point the robot trundles off with a fanfare (Figure 3) and the game is over.

Children played the tablet for three rounds of the game in a session, each self recorded through placing a tangible robot in a container. The game was demonstrated to the pupil and instructions followed a least-to-more prompting strategy (Doyle et al., 1988). At the start of the session they were reminded to choose the one that is more, and this was repeated after a first mistake, if this was immediately followed by a second mistake, a point was given to both arrays, and the pupil was instructed both verbally and through gesture to look at both, if this was immediately followed by a third mistake the researcher pointed to the correct array stating “more.”

Baseline and follow-up card game

Children were invited to identify which “had more spots” when presented with two cards. Pairs were presented in random order depicting small, medium, and large quantities in the ratios of 1:3; 1:2; 2:3; 3:4 and 5:6.

Transfer probe board activity

Children's performance was also compared to a non-digital board game that mirrored some aspects of the tablet game. It required children to choose between two groups of counters in ratios 1:3; 1:2; 2:3; 3:4 and 5:6. They then moved each of their chosen counters to form steps on a path (see Figure 4), and the other array was assigned to the researcher who moved her pile to form a parallel path. At each corner of the board game was a robot which was collected by the first player to reach that corner. This activity was presented on the session following the two sets of four intervention days.

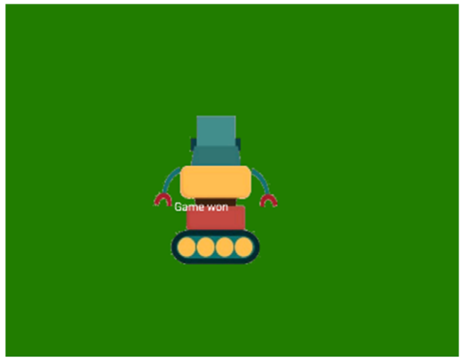

FIGURE 3 An assembled robot

Digital supplements

The second set of intervention days involved some further individual tailoring based on the responses in the first set. As the competition provided in the board game had proved motivating to the pupils, the children were given a two player version of the tablet game, identical to the first but with one of the researchers taking a second path and collecting pieces for a second robot. Two further alternatives were subsequently introduced, using a game Millie Moreorless that had greater variation in the ratios presented, and described more fully elsewhere (Porter, 2018). It required the selection between two arrays in order for Millie to move along the path to her final goal. Depending on the progress of the pupil they were either given a game involving the random presentation of ratios of 1:4; or 1:2 or 2:3. The latter involved selecting which was less, thus providing an extension to the response requirements.

One pupil who was failing to make progress using digital interventions was introduced to a series of non-digital games using coins, counters and cards in arrays of 1:2, and 1:3. Researcher 1 presented closed fists hiding one and two counters, or one and three counters for the pupil and researcher 2 to choose from and the pupil had to decide who had chosen the hand with more. Counters were put in two parallel lines to check the outcome. Researcher 1 presented the cards in a fan, face side away from the pupil and researcher 2, each of which chose one sight unseen, and on turning them over to reveal the stars, the pupil had to decide who had the card with more or if they had the same.

5.4 | Data analysis

Children's selection of response was recorded by the tablet for digital games and by two researchers for non-digital games. Trend data was generated to allow visual inspection of graphs in order to inspect within and between time and probe response patterns (Kourea & Lo, 2016). The standard mean difference was calculated between baseline/time 1, time 7 (post test) and the delayed post test to reveal positive and negative changes in performance. Field notes were hand written during each session, noting children's engagement with the activity and its component parts, the need for full prompts, the strategies adopted and any other notable occurrences. These were re-read against the trend data to examine contextual issues for changes in performance and to remove prompted performances from the scores.

6 | RESULTS

Looking across the eight pupils: comparing time 1 and time 7 reveals that five pupils (Ginny, Fred, Frank, Adi and Jonnie) made gains and that these were sustained or improved at the time of the delayed post-test (DPT) (see Table 3). Two pupils were performing well at the start of the programme, and either their performance was identical at DPT (Kat) or not sustained (Jo). The third pupil (Rizzi) continued to perform at the level of chance making no progress with digital materials.

The way the Robot game was set up the pupils did fewer trials with ratios 1:3 so we also looked at progress with different ratios over time. The majority of the pupils performed better on ratios of 1:3 than 1:2, with Jo and Kat attaining 100% on 1:3 from the start of the programme. Jonnie's profile is a notable exception with better scores on ratios of 1:2 at the outset which then tailed off.

The individual graphs (Figure 5) show the individuality of pupils' routes to progress together with their variability in responding, with few graphs showing a clear progression followed by a period of stability. If practice led to perfect performance we might expect the lowest scores to be at time 1 and the highest scores to be at post test.

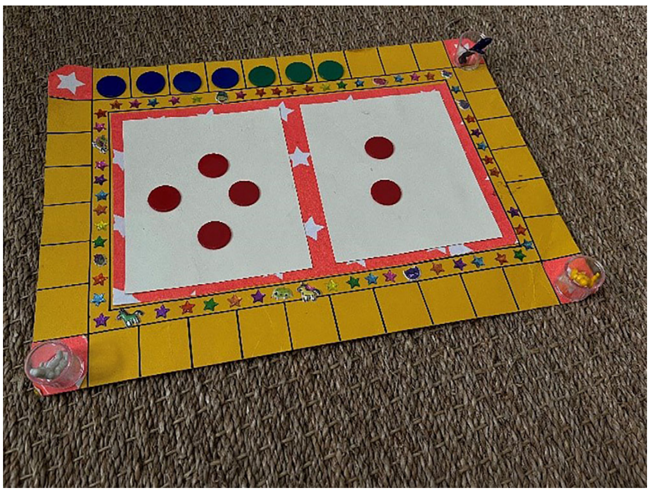

FIGURE 4 The nondigital board game

6.1 | Changes in performance on nondigital non-taught activities

6.1.1 | Card game

We looked at the mean change in performance on the card game (Table 4) from base-line through probe 1 (at time 4) and probe 2 (at time 7) and the delayed post test for the taught ratios of 1:2 and 1:3 and the untaught ratios of 2:3, 3:4, and 5:6. Four pupils showed changes in taught examples on the second probe (Ginny, Rizzi, Kat and Adi), but only two of these (Kat and Adi) also improved performance on the non-taught examples. Four (Ginny, Kat, Adi and Jonnie) also demonstrated positive changes at the time of the delayed post-test. It should be noted that there were also changes in the negative direction for two pupils (Frank and Jonnie) on the second probe for taught examples, one of whom, Frank continued to have a decrease in performance at the time of the delayed post-test.

6.1.2 | Non-digital board game

Data from pupils' performance on the board game reveals that four pupils (Ginny, Rizzi, Kat and Adi) showed positive changes, ranging

TABLE 3 Comparison of mean scores on percentage of correct digital trials

	Time1	Time4	Time7	Difference T1–T7	Delayed post test	Difference T1–DPT
Ginny	72	85	86	14	100	28
Fred	82	96	100	18	100	18
Jo	100	92	100	0	88	−18
Rizzi	56	48	50	−6	56	0
Frank	67	85	94	27	95	28
Kat	92	89	83	−9	92	0
Adi	80	84	95	15	96	16
Jonnie	55	57	85	30	91	36

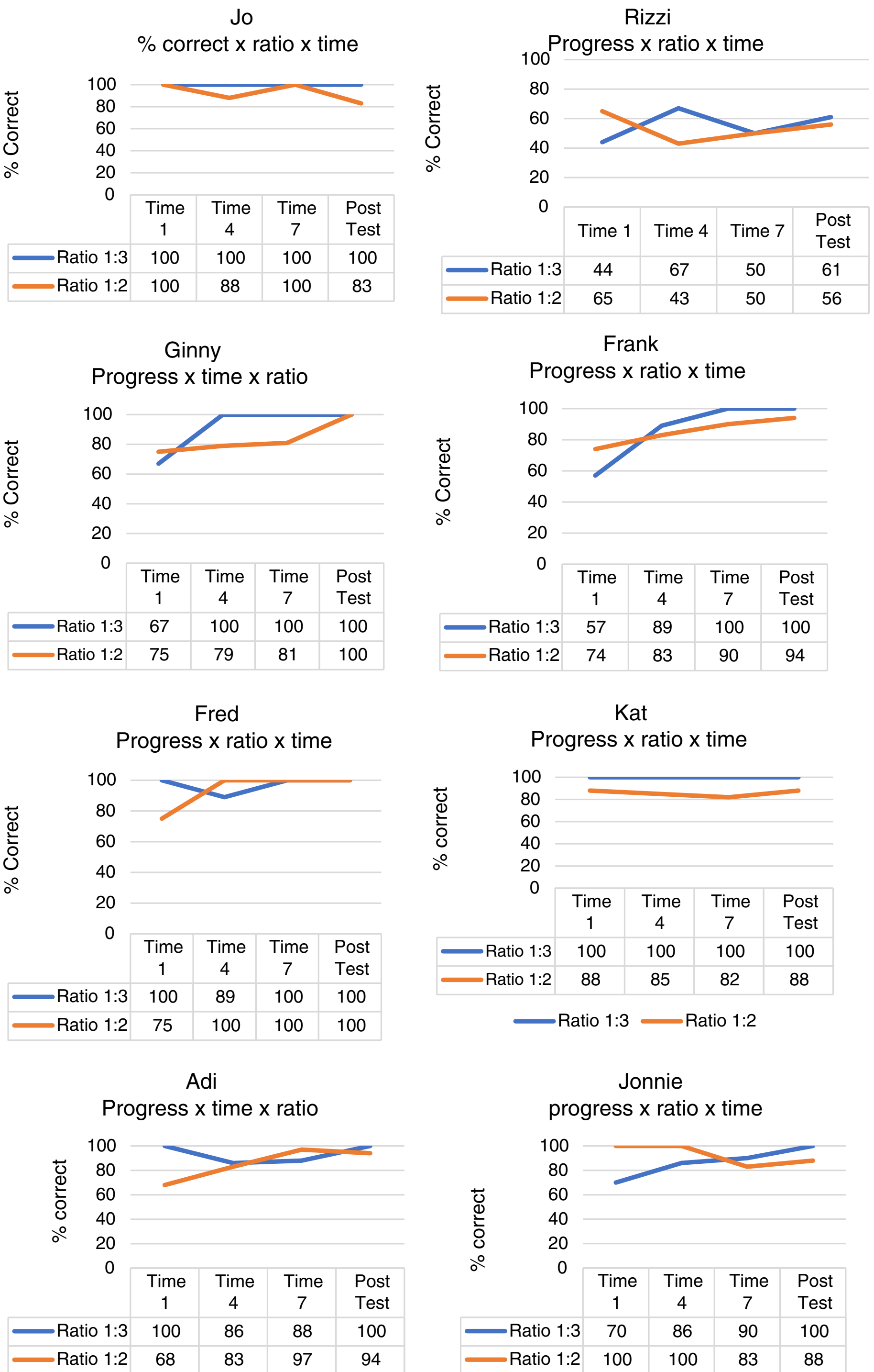

FIGURE 5 Individual pupils responses to the digital game by time and ratio

from 17–40% from probe 1 to probe 2 and this was true for both taught ratios and untaught ratios. The higher performing three pupils made no gains – in two (Fred and Jo) instances because they were at a ceiling level.

Looking across both the nondigital tasks we can see that on the test ratios of 1:3 and 1:2, two children performed at ceiling level, four had improved their scores at post test and four had scores above their baseline at delayed post test. Again however we can see some noise in the data- with Franks scores decreasing over time.

Turning to the board game, with the exception of Jonnie- all children improved their performance on the board game or continued to perform at ceiling level (see Table 5). One analysis of the difference between progress on the board game and the card game is that scores were lower on the latter. Children's comments as they played the board game suggested they were more attentive because of the social and competitive element.

TABLE 4 Card game: Mean changes on taught ratio 1:3 and 1:2 (all ratios, taught and untaught)

Name	Baseline (BL) cards	Probe 1	Probe 2	Difference BL–P2	Delayed post test (DPT) cards	Difference BL-DPT
Ginny	83 (87)	33 (67)	100 (87)	17 (0)	100 (93)	17 (6)
Fred	100 (93)	83 (87)	100 (87)	0 (−6)	100 (100)	0 (7)
Jo	100 (100)	100 (87)	100 (75)	0 (−25)	100 (87)	0 (−13)
Rizzi	83 (67)	67 (53)	100 (67)	17 (0)	83 (67)	0 (0)
Frank	83 (69)	86 (81)	71 (53)	−12 (−16)	67 (74)	−16 (5)
Kat	86 (76)	86 (94)	100 (87)	14 (11)	100 (100)	14 (24)
Adi	83 (63)	100 (80)	100 (94)	17 (31)	86 (94)	3 (31)
Jonnie	60 (67)	83 (67)	44 (59)	−16 (−8)	86 (75)	26 (15)

TABLE 5 Board game probes: Mean changes 1:2 and 1:3 (all ratios, taught and untaught)

Name	Probe 1 board game	Probe 2 board game	Difference P1–P2
Ginny	60 (75)	100 (90)	40 (15)
Fred	100 (91)	100 (89)	0 (−2)
Jo	100 (100)	100 (100)	0
Rizzi	43 (46)	100 (90)	57 (44)
Frank	80 (92)	80 (73)	0 (−17)
Kat	63 (50)	80 (64)	17 (14)
Adi	63 (50)	100 (90)	37 (40)
Jonnie	60 (70)	50 (80)	−10 (10)

6.1.3 | Two-player performance

To look further at the role of competition we compared playing the digital game as a single player and with a partner. The two-player version was introduced after the first four intervention sessions and in order to look at comparable data across the group we examined performance across contiguous sessions (usually session 5 and 6). Five children (Ginny, Fred, Jo, Frank, Jonnie) performed better with the two-player version, although in most instances the difference was modest (see Figure 6). The exception was Jonnie, for whom the introduction of the two player appeared to be a turning point. His pattern of engaging on the 1player was to go too fast and then make mistakes and this upset him. The introduction of a second player focused his mind, "you loser" said to partner when he won. A similar competitive response was made by Fred, whose profile reveals that the games with two people, both digital and nondigital, version focused his attention. Our field notes revealed him saying, "I'm going to win!" "Yes" said pulling in clenched fists. On several occasions he prompted his partner to make the wrong choice, "go for 4" and when she does he laughs. Frank also improved his game. Notably his best performance- 100% correct with both ratios was achieved when we introduced the two player version- scores he then achieved on two further separate occasions. Our field notes also revealed that this achievement was made against a backdrop of people arriving in school, giving weight to the view that at least for some students playing with another person focuses attention. Kat and Adi however found the introduction of a second player a distraction. Adi's initial

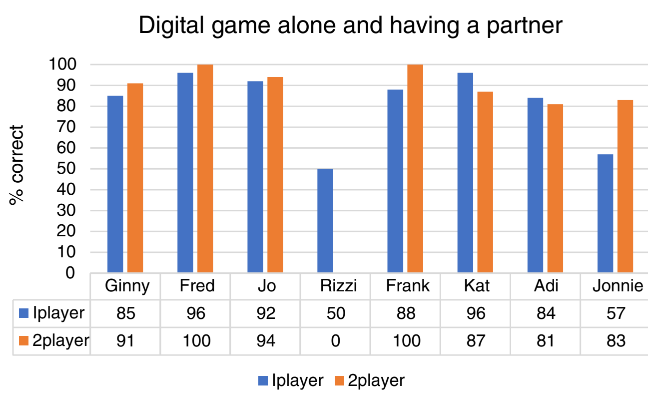

FIGURE 6 Comparison of playing alone or having a partner on the digital game

performance on the board game was poor- as if the introduction of a partner gave her too much to think about, as it prompted social interaction and helping her partner to choose the one with more, making it a co-operative game.

7 | DISCUSSION

In order to test out our bespoke digital game to improve pupils' performance on identifying which is more when presented with choices between arrays of ratios 1:3 and 1:2 we examined whether there was progress in the form of increased accuracy of identification of "more" between time 1 and time 7 and whether this was sustained for the delayed post-test (RQ1). Using a probe design we examined whether this was reflected in changes in performance with the same ratios on two non-digital tasks, a card game, and a board game that mirrored more closely the conditions of the tablet game. We also examined whether these changes were reflected in changes in performance on more demanding ratios (2:3. 3:4 and 5:6) (RQ2). We report here firstly at a group level, and then an individual level, drawing on qualitative data to explore the "noise" in the data.

Using aggregated mean scores for performance of 1:3 and 1:2 ratios, five of the pupils improved in their performance on the digital games and this was sustained or improved at the time of the delayed post-test. In most cases, but not all, the changes were of a modest nature. Two further pupils performed near or at ceiling at time 1, and made no sustained gains or their performance declined. One pupil continued throughout to perform at the level of chance on these digital tasks.

Looking at whether exposure on the digital games was reflected in improved performance on the card game, four pupils revealed improvements between baseline and probe 2 (time 7) and four between baseline and the delayed post test and this was true for both taught and untaught examples. Performance on the nondigital board game revealed improvements for all but one pupil between probe 1 and probe 2 or sustained ceiling level performance. In fact at probe 2 five of the pupils performed at ceiling level on the board game on the taught ratios and two further pupil performed at 80% correct. Five pupils also improved their performance for non-taught examples.

Data from the nondigital board game led us to compare performance on the single player version of Anna's robot with the two-player version. Like the non-digital version of the board game this context revealed increased levels of engagement for five of the pupils, at a modest level for some whose performance was already high but notable for others whose performance was variable.

In answering our research aim of finding out whether pupils with DS improve their understanding of magnitude, it would be however premature to conclude that the digital game is effective in bringing about change through examining the data only at the level of the group. Individual level data illustrates the variable level of responding. If we take the five pupils who demonstrated some improvement in performance across the course of the intervention, only two of these revealed improved performance on the digital probes of the card game, although they did so for both taught and untaught ratios, and those two also made increases in their performance on the non-digital board game. We can only therefore make a convincing case for the utility of the learning for two of the eight students.

We can also make a case for the ineffectiveness of this digital game as a learning tool for one pupil (Rizzi) who performed at the level of chance on the digital games. We were concerned that she was being exposed to high levels of failure (she scored 0 when we introduced her to the two player version). This led us to build on her relative strengths with tangible games in which she then scored 100% with ratios 1:3 and 100% with 1:2 on her final activity. She demonstrated similar improvements on the board game at the time of the second probe (100% correct for both ratios) and for the card game although this was not sustained at the delayed post-test.

For one pupil, Jo there was relatively little margin for improvement until we introduced an extension activity. His individual profile revealed that he was 100% correct across the duration for the easier ratio of 1:3 and more than 80% correct for ratio 1:2. He was already 100% correct for the taught ratios on the cards and similarly on the non-digital board game. (The probes therefore only functioned in relation to the non-taught examples). The task for him was therefore about sustaining attention. He got angry with himself when he made a mistake and the quick responses and feedback of the digital game drew his attention to this. We introduced an extension activity to provide a greater level of challenge but, mindful of his response to making errors, we gave him a choice between that and the easier game. The harder task involved discriminations of ratio 2:3 and identifying *less* rather than more. He progressed between scores of 76% correct to 97%. Thus, while he was limited in making progress in the group task, he demonstrated greater levels of understanding on his harder individualized task.

7.1 | The digital fit with the learning task

A close analysis at an individual level provided new insights into the task and its fit with technology. The game was designed to develop better awareness of numerosity, without requiring an ability to count correctly, an area of difficulty for many pupils with DS, our group included. The game required sustained attention, scanning two arrays, and discrimination of which is more. Individual graph profiles reveal some instability in responding with peaks in their performance at different times. Others have noted disengagement when skills have been accomplished (Wishart & Duffy, 1990) and in the number race when bored (Sella et al., 2021; Wilson et al., 2006). Serious games therefore have to hit the optimum level of challenge together with engaging sustained attention. For the majority of pupils, but not all, the introduction of a two-player version appeared to focus the mind, and bring about changes in attention and motivation. It introduced some novelty while not ostensibly changing the task demands. For those pupils who were motivated to win, it prompted them to use their awareness to predict, to strategize, and in some cases to try and cause their opponent to lose. The makers of the number race introduced new characters to offset boredom (Wilson et al., 2006), here we introduced the choice of a new game. There is some evidence from studies of young children with DS of a preference for novel objects (Onnivello et al., 2019) but also cautions that this can be distracting (Porter, 2018; Gulliford et al., 2021), suggesting that the introduction of novelty needs to be mediated by teachers and monitored for its impact on learning. Pupils did not respond in the same way to the introduction of change, for some it depressed performance. Game makers need to routinely introduce more choice (Valenza et al., 2019) while keeping the level of difficulty consistent, enabling learners with DS to consolidate their learning.

An important element of fostering sustained attention is the choice of response. The game deployed both a tap and drag response. Repeated tap responses can encourage pupils to rush, with limited prior scanning between alternatives, nor opportunity for prompting. The more effortful dragging response served to slow pupils down, encourage scanning and attention to the visual array. Our field notes revealed that although some pupils found dragging the spots onto the footpath difficult due to their size, all were sufficiently motivated to do so, and it reinforced the notion of a numberline (Siegler & Ramani, 2009), while also providing fine motor practice. During the development phase we introduced sound to indicate that the spot was in place offsetting potential frustration of completing the activity (Martin-Gutierrez et al., 2021) and providing a better fit between the task and the child's motor skills (Nacher et al., 2018).

The pupil who made no gains with the digital game but did progress with cards and coins illustrated the way in which different learning contexts provide different affordances. For her, the tangible materials provided more sensory information and the games prompted anticipation. It was also a more controllable teaching situation using full feedback (hesitation, speed of responding, scanning responses) to inform the next teaching presentation. For this pupil, at the earliest stages of awareness, non-digital contexts provided the speed and ease of tailored teaching that is not yet matched by digital equivalents (Valenza et al., 2019).

7.1.1 | Limitations

There are some limitations to this study, many reflecting the real world school context in which it was under-taken. Little comparable data exists to demonstrate the representativeness of this group of children although, with the exception of their oral skills, their counting skills are commensurate with the larger sample of children with DS drawn from mainstream schools by Nye et al. (1995). They also reveal the expected relative difficulty in relation to enumeration and production of sets (Onnivello et al., 2019). However, given that some pupils reached a ceiling level, our sample could have been increased. We invited all pupils with Down syndrome in the school between 7 and 14 years to take part and parents of eight of the ten children responded to the invitation. Further studies could be based on samples from additional schools.

A longer intervention period might have led to more learning (Sella et al., 2021) but was limited by access to pupils. While three visits a week were planned, all but two pupils had absences, due to ill health, day trips out, or sports fixtures and their sessions were re-scheduled. We planned a more extended time gap between the intervention and follow-up. Sports days, residential visits and end of term celebrations, necessitated an earlier return visit.

Technical issues meant that the planned advanced levels of Anna Robot game were not available at the point of the intervention that they were required. This meant that some pupils reached a ceiling level and an alternative digital game was introduced. This unplanned set of circumstances had some advantages as it served to provide novelty and enhance engagement in the young people. It also provided new opportunities for individualizing access to the games thereby offsetting the likelihood that playing the same game three times a week can lower thresholds of attention.

7.1.2 | Implications

Our findings have implications for evaluation methodology. Our approach was driven by two aspects; namely that children should act as their own controls, allowing us to look within as well as across the data, representing variation within the group. Secondly, the possibility of introducing additional individualisation to the intervention programme to enable a better fit to pupils' strengths and needs. The aim was therefore to test out the effectiveness in bringing about new learning through a number of bespoke features through an ethical methodology for evaluation.

Evaluating digital games requires an approach that maximizes pupils' engagement and interest in the activity in an ethical way. It avoids exposure of pupils to repeated failure and/or frustration, by not adopting test conditions that require the presentation of extended baselines. It is inextricably linked to decisions around the length and duration of the intervention, the point at which sessions stop. It is also reflected in decisions about inclusion criteria and the extent to which they reflect full variation within the group. An ethical methodology emphasizes the importance of pupils actively opting in to the research rather than it being timetabled as an additional lesson; of them knowing what the expectations are in terms of duration, so that their informed consent is asked for the specifics of the activities rather than soliciting a more generalized agreement to take part. It therefore beholds the researcher to ensure that the activities are engaging and that additional recognition for effort is provided. In this respect the use of individual sticker booklets marking numbers of sessions completed were integral to the session, as were the use of robots and their spaceship pots to mark the number of games to be played. There are also important negotiations to be had in school about the timing of activities to ensure that pupils do not miss a favoured activity- and in the context of special schools this may not have a fixed timing but vary from day to day.

The study revealed the importance of disaggregating data to examine individual profiles rather than only reporting on group data and the challenges of making sense of the unstable responding and the role of qualitative data in doing so. As Ledford et al. (2018) point out high variability interferes with the ability to identify a functional relationship. The introduction of a variation of the programme and extension activities served to provide some novelty and keep the pupils engaged. This again has implications for the design of the study. The inclusion of non-digital probe sessions contributed to variety. Our data revealed that the choice of probes provides space to compare responses in alternative contexts. Perhaps surprisingly if we look at specific individuals, progress on the probes did not necessarily mirror progress on the digital task. Frank for example made progress in the digital context but this was not reflected in improved scores on the non-digital tasks. Kat made no progress in the digital setting but did on the non-digital tasks. A longer intervention period would have enabled us to assess whether this reflected their different entry level competence on the task or their preference for particular learning contexts.

7.1.3 | Design and intervention implications

Digital games can provide a medium for promoting attention to differences in magnitude for children with DS but there is variation in their utility despite a bespoke design. Here we deployed two approaches to being bespoke. The game itself was designed with the difficulties of children with DS in mind. This included the specifics of the interface, clarity of the visual presentation, the fine motor requirements, the use of auditory feedback, the pace of learning (children may take longer to respond) the analysis of progression in the target skills and provision of small steps, and the automatic adjustment in the level of difficulty. We added further individualization through the provision of different extension activities, which in one instance was nondigital. Arguably with hindsight some of these extensions could have been built into the game. However that does not remove the need for a mediator, as Gulliford et al. (2021) argue "Pedagogical facilitation… must be available in order to realize technology's potential benefits" p. 102 particularly in the context of learners with special educational needs.

The benefits can be considered in two ways. For the specific task here, five out of eight pupils improved performance and this was sustained at the delayed post-test. Two pupils generalized their learning to untaught non-digital settings. More general learning to learn skills also occurred. Playing a serious game promotes sustained attention and engagement. The use of a "least to most" prompting regime encourages the child to slow down and look at all of the display. Success builds confidence and encourages children to be less impulsive (Gulliford et al., 2021). These are important skills for interacting with all digital material.

This study also draws attention to the individual nature of the affordances provided by different contexts. This was well illustrated when the game included a second player. While this focused the attention, raised motivation and improved performance for some it proved inhibiting and distracting for others. The meaning of the activity changed. This has implications for the teacher in drawing flexibly on different learning tools based around making a fine grain analyses of pupils' responses in different contexts. The aim is to provide the optimal level of challenge, one that is slightly above their current level, that they can achieve with support. This requires an analysis not just of how many pupils get right, but which ones they succeed on, what type of response is required, and what level of support is necessary for them to be successful.

Through looking at individual profiles, we have evidenced the variability of responding, with some pupils showing negative "gains" at some points of the intervention. The speed of making a digital response can overtake a pedagogic prompt with implications for the deployment of different touch movements. For example, it was easier to prompt when a drag and drop motion was required than when a simple tap was needed. Non-digital settings provide an easier setting for supporting the child, and may go towards explaining why the child with the greatest difficulty in this area only made progress using tangible materials. Further evidence is needed of the relationship between performances in different contexts through studies over a longer period of time with pupils bringing a range of different strengths and needs.

8 | CONCLUSION

This study contributes to a limited field of research on the development of bespoke ICT for young people with intellectual disability and its impact Shahid et al. (2021). It illustrates some of the challenges and complexities of both design and evaluation research. With respect to design, games have to hit the optimum level of challenge and pace while also sustaining "minds on" engagement in order to consolidate learning. This study indicates that the careful introduction of novelty, seen here through the introduction of a second game player, or change of character and narrative can help to focus attention on a key skill that underpins an understanding of number.

With respect to evaluation, the methodology of single case studies revealed the noise in the data which can so easily be camouflaged when group responses only are reported. It illustrated both the variability between pupils with Down syndrome but also the instability of responses with daily fluctuations contributing to an uneven path of progression. From the perspective of the teacher, single case studies can provide greater ecological validity, mirroring more closely their own classroom experience. They also promote an ethical approach to evaluation, facilitating the tailoring of teaching to pupil needs, rather than a more generic approach whereby some individuals may experience repeated failure, reducing self-confidence and the motivation to succeed.

The study had advantages over other intervention studies in that it met the concerns of others (e.g., Sella et al., 2021) in targeting a particular understanding and it tested out the impact on other tasks. Although the gains were modest, learning through bespoke serious games shows promise but they do not replace the need for expert support (Sella et al., 2021), nor for access to a range of teaching contexts. We can therefore understand "bespoke" properties as firstly comprising a design that is tailored to the strengths and needs of the particular group, and secondly requiring the teacher to make a fine grain analysis of its place for the individual learner amongst a range of learning contexts.

ACKNOWLEDGMENTS

Anna's Robot was designed by Anna Goclowska with supervision from Professor Richard Mitchell at the Department of Computer Science, University of Reading.

CONFLICT OF INTEREST

The author declares no potential conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jcal.12685.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Jill Porter https://orcid.org/0000-0002-8507-6260

REFERENCES

Abreu-Mendoza, R. A., & Arias-Trejo, N. (2015). Numerical and area comparison abilities in Down syndrome. *Research in Developmental Disabilities*, *41-42*, 58–65. https://doi.org/10.1016/j.ridd.2015.05.008

Agheana, V., & Duta, N. (2015). Achievements of numeracy abilities to children with Down syndrome: Psycho-pedagogical implications. *Procedia-Social and Behavioural Sciences*, *186*, 38–45. https://doi.org/10.1016/j.sbspro.2015.04.068

Ahmad, W. F., Muddin, H. N., & Shafie, A. (2014). Number skills mobile application for Down syndrome children. In *International Conference on Computer and Information Sciences (ICCOINS), 2014* (pp. 1–6). IEEE. https://doi.org/10.1109/ICCOINS.2014.6868844

Baddeley, A.D., & Jarrold, C. (2007). Working memory and Down syndrome. *Journal of Intellectual Disability Research*, *51*(12), 925–931.

Belacchi, C., Passolunghi, M. C., Brentan, E., Dante, A., Persi, L., & Cornoldi, C. (2014). Approximate additions and working memory in individuals with Down syndrome. *Research in Developmental Disabilities*, *35*(5), 1027–1035.

Browder, D. M., Trela, K., Courtade, G. R., Jimenez, B. A., Knight, V., & Flowers, C. (2012). Teaching mathematics and science standards to students with moderate and severe developmental disabilities. *The Journal of Special Education*, *46*, 26–35. https://doi.org/10.1177/0022466910369942

Clarke, B., & Faragher, R. (2014). Developing early number concepts for children with Down syndrome. Chapter 6. In R. Faragher & B. Clarke (Eds.), *Educating learners with Down syndrome* (pp. 146–162). Routledge.

Cregg, M., Woodhouse, J. M., Pakeman, V. H., Saunders, K. J., Gunter, H. L., Parker, M., Fraser, W. I., & Sastry, P. (2001). Accommodation and refractive error in children with Down syndrome: Cross-sectional and longitudinal studies. *Investigative Ophthalmology & Visual Science*, *42*, 55–63.

Doyle, P. M., Wolery, M., Ault, M. J., & Gast, D. L. (1988). System of least prompts: A literature review of procedural parameters. *JASH*, *13*(1), 28–40. https://doi.org/10.1177/154079698801300104

Felix, V. G., Mena, L. J., Ostos, R., & Maestre, G. E. (2016). A pilot study of the use of emerging computer technologies to improve the effectiveness of reading and writing therapies in children with Down syndrome. *British Journal of Educational Technology*, *48*, 611–624. https://doi.org/10.1111/bjet.12426

Feng, J., Lazar, J., Kumin, L., & Ozok, A. (2010). Computer usage by children with Down syndrome: Challenges and future research. *ACM Transactions on Accessible Computing*, *2*(3), 1–44. https://doi.org/10.1145/1714458.1714460

Fidler, S. H., & Rogers, S. (2006). Early learning and adaptive behavior in toddlers with Down syndrome: Evidence for an emerging behavioural phenotype? *Down Syndrome, Research and Practice*, *9*, 37–44. https://doi.org/10.3104/reports.297

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.

Hasan, P. J., & Messer, D. J. (1997). Stability or instability in early cognitive abilities in children with Down's syndrome? *The British Journal of Development Disabilities*, *43*, 93–107. https://doi.org/10.1179/bjdd.1997.011

Horner, R. D., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, *11*, 189–196. https://doi.org/10.1901/jaba.1978.11-189

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S. L., & Wolery, M. (2005). The use of single subject research to identify evidenced-based practice in special education. *Exceptional Children*, *37*, 165–179. https://doi.org/10.1177/001440290507100203

Karmiloff-Smith, A., Al-Janabi, T., D'Souza, H., Groet, J., Massand, E., Mok, K., Startin, C., Fisher, E., Hardy, J., Nizetic, D., Tybulewicz, V., & Strydom, A. (2016). The importance of understanding individual differences in Down syndrome. *F1000Research*, *5*, F1000 Faculty Rev-389. https://doi.org/10.12688/f1000research.7506.1

King, S. A., Powell, S. R., Lemons, C. J., & Davidson, K. A. (2017). Comparison of mathematics performance of children and adolescents with and without Down syndrome. *Education and Training in Autism and Developmental Disabilities*, *52*(2), 208–222.

Kourea, L., & Lo, Y. (2016). The educational validity and utility of single case design research in building evidence-based practices in education. *International Journal of Research & Method in Education*, *39*(4), 349–364. https://doi.org/10.1080/1743727X.2016.1160278

Ledford J. R, Zimmerman, K.N., Schwartz I & Odom, S. (2018) *Guide for the use of single case design research evidence from the division for research of the council for exceptional children*. CEC-Divsion Research. https://cecdr.org/sites/default/files/2021-01/CEC-DR_SCD_Guide.pdf

Lemons, C. J., Powell, S. R., King, S. A., & Davidson, K. A. (2015). Mathematics interventions for children and adolescents with Down syndrome: A research synthesis. *Journal of Intellectual Disability Research*, *59*, 767–783. https://doi.org/10.1111/jir.12188

Martin-Gutierrez, J., & Del Rio Guerra, M. S. (2021). Analysing touchscreen gestures: A study based on individuals with Down syndrome centred on design for all. *Sensors*, *21*, 1328. https://doi.org/10.3390/s21041328

Monari Martinez, E., & Pellegrini, K. (2010). Algebra and problem-solving in Down syndrome: A study with 15 teenagers. *European Journal of Special Needs Education*, *25*, 13–29. https://doi.org/10.1080/08856250903450814

Nacher, V., Cáliz, D., Jaen, J., & Martínez, L. (2018). Examining the usability of touch screen gestures for children with Down syndrome. *Interacting with Computers*, *30*, 258–272. https://doi.org/10.1093/iwc/iwy011

Nascimento, L. S., Zagaro, N., & Martins L. B. (2020). Challenges of developing a mobile game for children with Down syndrome to test gestural Interface. *Information*, *11*(3), 159. https://doi.org/10.3390/info11030159

Nye, J., Clibbens, J., & Bird, G. (1995). Numerical ability, general ability and language in children with Down syndrome. *Down Syndrome Research and Practice*, *3*(3), 92–102.

Onnivello, S., Lanfranchi, S., & Zorzi, M. (2019). Chapter Eight – Mathematical abilities in Down syndrome. *International Review of Research in Developmental Disabilities*, *56*, 257–291.

Ortega-Tudela, J. M., & Gómez-Arizaw, C. J. (2006). Computer-assisted teaching and mathematical learning in Down syndrome children. *Journal of Computer Assisted Learning*, *22*, 298–307. https://doi.org/10.1111/j.1365-2729.2006.00179.x

Park, A. H., Matt, M. D., Wilson, A., Paul, M. D., Stevens, T., Harward, R., & Hohler, N. (2012). Identification of hearing loss in pediatric patients with Down syndrome. *Otolaryngology–Head and Neck Surgery*, *146*, 135–140.

Porter, J. (2014). Research and Pupil Voice. In Florian L. (Ed.), *Handbook of Special Education 2nd edn*, (Vol. 2, pp. 405–420). London: Sage.

Porter, J. (2016). Introducing magnitude representation with Millie Moreorless, a digital game for children with Down syndrome. *SLD Experience*, *75*, 21–27.

Porter, J. (2018). Entering aladdin's cave: Developing an app for children with down syndrome. *Journal of Computer Assisted Learning*, *34*(4), 429–439.

Porter, J. (2019). Discriminating quantity: New starting points for children with down syndrome about number. *International Journal of Disability Development and Education*, *66*(2), 133–150.

Sella, F., Lanfranchi, S., & Zorzi, M. (2013). Enumeration skills in Down syndrome. *Research in Developmental Disabilities*, *34*, 3798–3806. https://doi.org/10.1016/j.ridd.2013.07.038

Sella, F., Onnivello, S., Lunardon, M., Lanfranchi, S., & Zorzi, M. (2021). Training basic numerical skills in children with Down syndrome using the computerized game "the number race". *Scientific Reports*, *11*, 2087. https://doi.org/10.1038/s41598-020-78801-5

Shafie, A., Wan Ahmad, W. F., Mohd, N., Barnachea, J. J., Taha, M. F., & Yusuff, R. L. (2013). "SynMax": A mathematics application tool for Down syndrome children. In H. B. Zaman, P. Robinson, P. Olivier, T. K. Shih, & S. Velastin (Eds.), *Advances in visual informatics. IVIC 2013. Lecture notes in computer science, 8237*. Springer. https://doi.org/10.1007/978-3-319-02958-0_56

Shahid, M. I., Lai-Chong Law, E., & Verdezoto, N. (2021). Technology-enhanced support for children with Down syndrome: A systematic literature review. *International Journal of Child-Computer Interaction*, *31*, 100340. https://doi.org/10.1016/j.ijcci.2021.100340

Siegler, R. S., & Lortie-Forgues, H. (2014). An integrative theory of numerical development. *Child Development Perspectives*, *8*(3), 144–150. https://doi.org/10.1111/cdep.12077

Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—But not circular ones—Improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*(3), 545–560. https://doi.org/10.1037/a0014239

Spooner, F., Root, J. R., Saunders, A. F., & Browder, D. M. (2019). An updated evidence-based practice review on teaching mathematics to students with moderate and severe developmental disabilities. *Remedial and Special Education*, *40*, 150–165. https://doi.org/10.1177/0741932517751055

Tsikinas, S., & Xinogalos, S. (2019). Studying the effects of computer serious games on people with intellectual disabilities or autism spectrum disorder: A systematic literature review. *Journal of Computer Assisted Learning*, *35*, 61–73. https://doi.org/10.1111/jcal.12311

Valenza, M. V., Gasparini, I., & Hounsell, M. d. S. (2019). Serious game design for children: A set of guidelines and their validation. *Journal of Educational Technology & Society*, *22*(3), 19–31. https://www.jstor.org/stable/26896707

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "the number race", an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions*, *2*, 19. https://doi.org/10.1186/1744-9081-2-19

Wishart, J. (2005). Children with Down syndrome. Chapter 7. In A. Lewis & B. Norwich (Eds.), *Special teaching for special children?* Open University Press.

Wishart, J. G., & Duffy, L. (1990). Instability of performance on cognitive tasks in infants and young children with Down's syndrome. *British Journal of Educational Psychology*, *60*, 10–22. https://doi.org/10.1111/j.2044-8279.1990.tb00918.x

Yang, Y., Conners, F. A., & Merril, E. C. (2014). Visuo-spatial ability in individuals with Down syndrome: Is it really a strength? *Research in Developmental Disabilities*, *35*, 1473–1500.

How to cite this article: Porter, J. (2022). Evaluating performance on a bespoke maths game with children with Down syndrome. *Journal of Computer Assisted Learning*, 1–14. https://doi.org/10.1111/jcal.12685