

Number Express: A Digital Game to Improve Early Numeracy

Nicoletta Perini¹, Annamaria Porru², Korbinian Moeller^{1,3,4}, Tim Jay⁵, Daniela Lucangeli² and Francesco Sella¹

¹Department of Mathematics Education, Loughborough University, UK

²Department of Developmental Psychology and Socialisation, University of Padova, Italy

³Leibniz-Institut Fuer Wissensmedien, Tuebingen, Germany

⁴LEAD Graduate School and Research Network, University of Tuebingen, Germany

⁵School of Education, Faculty of Social Sciences, Nottingham University, UK

n.perini2@lboro.ac.uk

annamaria.porru@unipd.it

k.moeller@lboro.ac.uk

Tim.Jay@nottingham.ac.uk

daniela.lucangeli@unipd.it

f.sella@lboro.ac.uk

<https://doi.org/10.34190/ejel.22.10.3522>

An open access article under [CC Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)

Abstract: Early numeracy has repeatedly been observed to predict more advanced mathematical skills and vocational prospects. Accordingly, there is increasing interest in interventions using games and game-based activities to foster basic numerical skills. In the current study, we describe how we sought initial feedback on the game Number Express following a design-based research method and a study piloting the game with a group of four to six-year-old children to evaluate its usability and likability. The game consists of a train with six carriages for the player to fill in with numbers in the correct order. The game progresses in difficulty across several levels, moving from smaller to larger numbers and sequences in steps of 2, 5, and 10. Informative feedback is provided to assist players when they make incorrect responses. Associating numbers to carriages in the correct order earns the player points, with which they can buy items in a virtual shop, providing additional experience with calculating with money. During the design phase, we sought feedback from researchers, teachers, and an education consultant specialising in the learning and teaching of primary mathematics. We modified the game based on their suggestions. In addition, we ran a pilot study to test the game on a small group of preschool children (N = 45) evaluating its playability and user experience to allow for further development of the game. Overall, the results of the pilot study indicated that Number Express is straightforward to engage with, features easily memorable instructions, and incorporates an effective number line tool that supports number ordering skills. As a next step, we plan to use Number Express in a larger-scale intervention study to determine its effectiveness in improving children's early numeracy skills. In case the game proves to be effective, it has the potential to be a valuable tool for educators and parents in supporting their children's early numerical development.

Keywords: Digital game, Number ordering, Preschool and primary children, Early numeracy, Piloting

1. Background

Basic numerical skills are the foundation of more advanced mathematical skills (Watts et al., 2014) and significantly predict future mathematical learning and academic achievement (e.g., Atkinson et al., 2022; Duncan et al., 2020; Davis-Kean et al., 2022; Jordan, Hanich and Kaplan, 2003; Duncan et al., 2006). Around the age of two years, children typically start implementing counting routines to learn the cardinal meaning of number words, this means that “one” describes a set with one item, “two” a set with two items, etc. (Gelman and Gallistel, 1978; Sarnecka, 2015; Sarnecka, Goldman and Slusser, 2015). From there on, children progressively associate specific number words with specific numerosities in terms of symbol-magnitude associations. In this context, children also learn that number words reflect ordinal relations depending on their position within the counting sequence and children need to understand the ordinal relations among the corresponding digits, for instance, that 5 comes after 4 but before 6 (Sasanguie and Vos, 2018). The fact that numbers can be spatially represented in order on a number line (i.e., the position of a number provides information about its magnitude and ordinal relation as argued for cartesian axes, Sella et al., 2017) is another important characteristic of early numerical representations.

Not surprisingly, numerous interventions draw on games and activities to facilitate understanding of ordinal relations between numbers, but also cardinal number magnitude understanding by fostering children's skills in correctly locating numbers on a number line (e.g., Siegler and Ramani, 2008).

With the increasing availability of touch-screen devices such as smartphones and tablets, specific games and apps have been developed to support children's early numerical learning (Klawe, 1999; Annetta et al., 2009; Sella et al., 2016; Sella et al., 2021; Benavides-Varela et al., 2020).

Besides acquiring subject-specific mathematical skills and knowledge, such digital games also seem to increase students' motivation and promote positive attitudes towards learning (e.g., Ke, 2009; Pareto et al., 2011; Crittenden, Biel, and Lovely, 2019). For instance, The Number Race (<http://www.thenumberrace.com/nr/home.php>) is an example of a research-informed digital game (Wilson et al. 2006;) that has been evaluated in both typically and atypically developing children (Sella et al. 2016; Rasanen et al., 2009; Sella et al., 2021, Sella et al., 2016). The game presents numbers in various forms: as digits, spoken number words, and visualized quantities through sets of objects. The player is tasked with selecting the larger of two numbers, beginning with concrete sets and gradually advancing through spoken and written forms until only written numbers are used. Moreover, the tokens earned by the player are moved to a racetrack, illustrating how numbers are mapped to a number line. The game supports multiple numerical skills, such as comparing symbolic and non-symbolic numerical quantities, counting, and simple arithmetic. During the game, children are also asked to move their game character along a line, therefore reinforcing the ordinal components. Other digital games have directly used the concepts of the number line, where players are required to place numbers along a visual number line (for example Rescue calculatoris, a mental number line training for children with developmental dyscalculia, Kaser et al. 2013; Semideus, a game on fraction processing, Ninaus et al. 2017). Even though these games support different basic numerical skills, including the use of a number line, they do not specifically target number-ordering skills.

This seems surprising as there is accumulating evidence that number-ordering skills are crucial for the development of early numerical skills (Devlin et al. 2024; Morsanyi, O'Mahony and McCormack, 2017). Additionally, in a recent review by Balladares and colleagues (2023), the authors selected nineteen articles exploring the relationship between board games and mathematical skills. Interestingly, only one game was reported to focus on number ordering.

Conversely, there are plenty of apps and online games designed to help children learn numbers by focusing on their order, such as Bud's Number Garden (BITESIZE, 2024) or Jumpy: Counting by Ones From 1 to 100 (education.com, 2024). However, for most of these apps, the research underlying the game is not clear, and many are freely accessible only for a limited number of trials.

Therefore, we decided to create a freely available game facilitating number order processing.

Against this background, the current article describes the development of the app Number Express which is intended to facilitate early learning about numerical order by requiring children to place numbers in the correct order on a number line. In this study, we describe how we gathered initial feedback on the game using a design-based research method. Moreover, considering that the use of digital games with kindergarten children could present challenges—such as understanding the rules and instructions—we also conducted a pilot study with a group of four- to six-year-old children to evaluate the game's usability and likability.

In the following, we will first introduce the main conceptual characteristics of the game and initial feedback on its usability by numerical cognition researchers and educational practitioners before we report the results of a pilot intervention study. Finally, we discuss next steps to further develop the app and evaluate its effectiveness in a larger-scale intervention study.

2. Conceptual Characteristics of the Game

2.1 The Theme and the Game's Main aim

The narrative of *Number Express* is situated at a train station where the train conductor has to renumber carriages of an incoming train because some carriage numbers have fallen off. The train has six carriages, some of which still have their number label on (serving as reference points for numbering the other carriages), while others do not (reflecting target numbers to be allocated by players).

When starting the game, children first choose between different animal characters representing the train conductor they want to play with and type in their name (see Figure 1).

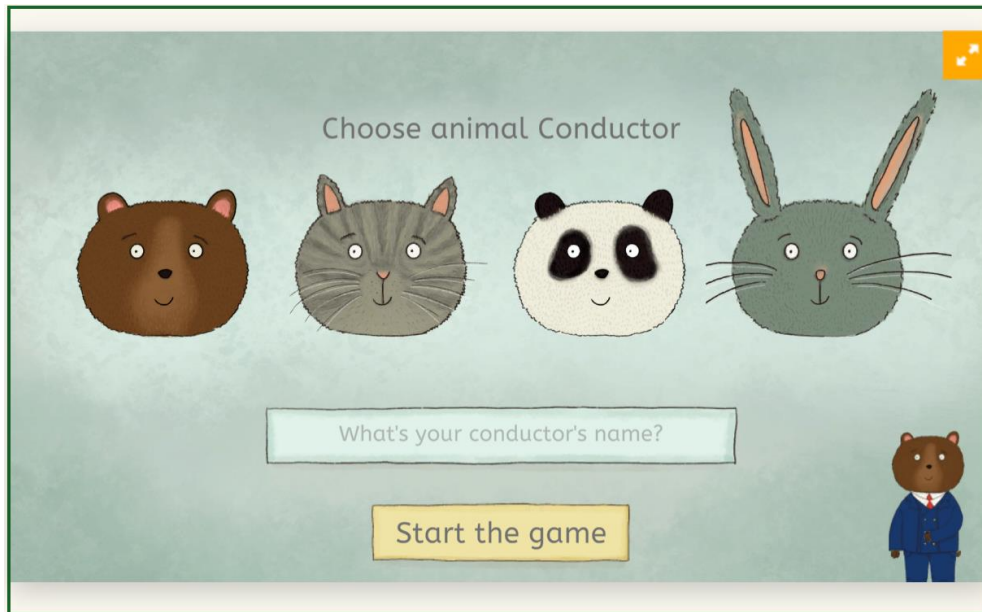


Figure 1: Starting screen to choose the character of the train conductor

Then, children have to select the level they want to start playing from (see Figure 2).

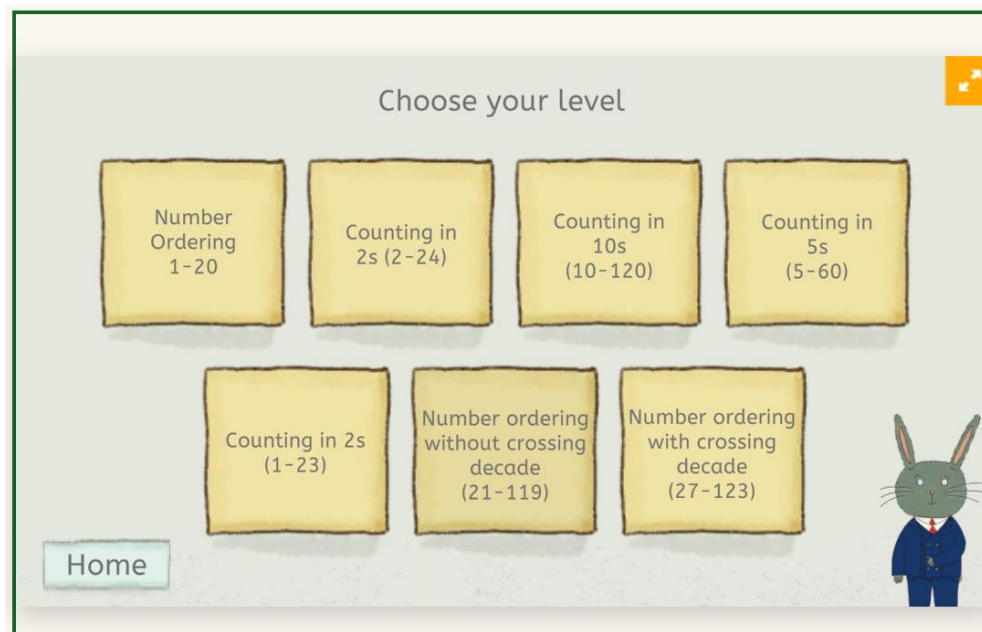


Figure 2: Starting screen to choose the level the player wants to play in

Then, a train arrives on the screen, and a target number (to be associated with one of the unlabelled carriages) is shown at the top-left of the screen together with the sound of the respective number word emphasizing the association between visual and verbal representations of the respective number (e.g., “four”; and 4, cf. Figure 3). The player then has to drag the presented target number to the respective carriage, considering the given numbering on the other carriages as references or simply click on the chosen carriage.

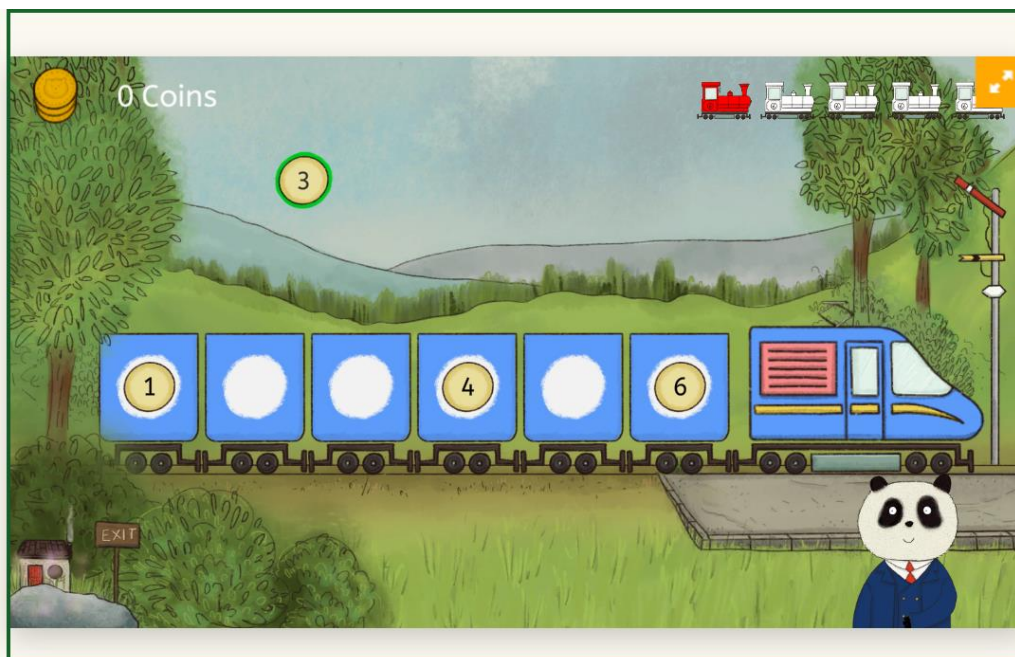


Figure 3: Main gameplay screen showing the train with already numbered carriages. The number 4 with a green circle around it is the number that the player has to put on the right carriage. At the top of the screen, the number of coins earned by the player is given (the number next to the coin's image) as well as the number of rounds a player has already completed within a level (the coloured train indicates that this is the first round played. When the player plays the second round the second train will turn orange and so on). Each level has five trains. The button *exits* (the sign on the bottom left of the screen) leads the player to exit the game. The button *home* (the little house depicted on the bottom left of the screen) leads the player to the game's homepage

After associating the first target number with a respective carriage, the next one will appear until all carriages are numbered. At that point, the locomotive shines, and children can click on it to make the train leave the station. In case target numbers were associated with carriages to reflect a correct order (e.g., 1, 2, 3, 4, 5, 6 in Figure 3), the train leaves the station, and another train appears on the screen. When some of the numbers are not put on the carriages in correct order, the train does not leave the station and the numbers which are on the wrong carriages will be moved to the top of the screen with informative feedback appearing on the screen. In case the player puts the numbers on the wrong carriages the first feedback that they receive is a visual and verbal message (Good try! But some of the numbers are not at the right place). Then they can try again to put the target numbers onto the right carriages. When they put the numbers in the wrong order again, they receive the same message as for the first reply. In addition, a number line appears right under the train so that the player can use the number line to copy the position of some numbers (See paragraph 2.4 and Figure 5 for more information on this game's informative feedback). For every train the player makes leave, they are awarded one coin as a virtual incentive. After five trains successfully leave the station, the player will have earned five coins to spend in a virtual shop (Figure 4) allowing them to buy accessories for the train (e.g., stickers to decorate the train or the possibility to change the shape of the train) and clothes or accessories for the train conductor (e.g., a new outfit, a hat, or a toy). To pay in the shop the player can only use the virtual coins earned playing the game. They can also change the train conductor character for free (e.g., panda, bear, cat, and rabbit). Once they are done with their shopping, they can return to playing the game.

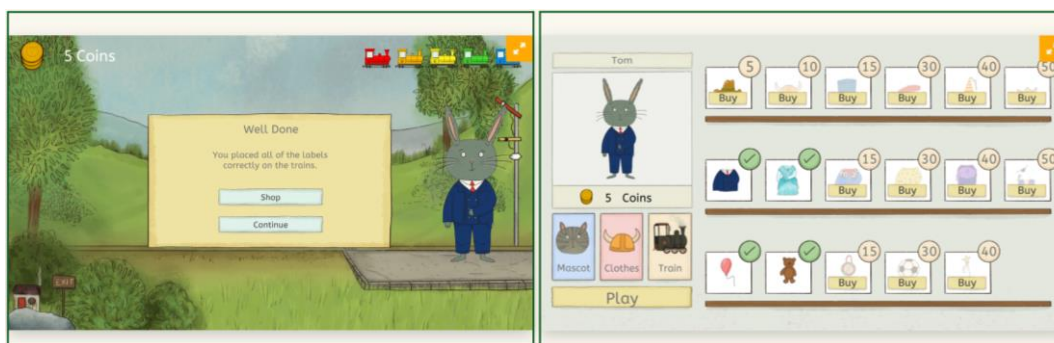


Figure 4: The shop. On the left panel, the screen is displayed on which the player can choose to go to the shop or continue playing after five rounds of playing. In the latter scenario, the player will continue accumulating the coins earned during the game. The right panel shows the screen of the shop. The player can see their train conductor with their name and the amount of virtual coins they can spend in the shop. Once they have completed their shopping the amount of coins will change accordingly. Clicking on the button with the label Clothes allows the player to see all the goods they can buy for the train conductor. Clicking the button with the mascot on allows the player to change the train conductor for free. Finally, clicking on the button with the train on allows the players to see all the accessories they can buy for the train

2.2 The Levels

The game comes with 118 levels of difficulty, which mainly vary depending on the number range and the respective number of reference numbers already on the carriages. For instance, level 1 only involves numbers between 1 and 6 with three reference numbers given, whereas level 60 involves numbers between 71 and 76 with only one reference number depicted on one of the carriages. Each level involves 5 rounds (5 trains with all the carriages labelled).

It is possible to start the game at seven different starting points (Table 1) for levels operating on the same numerical content (i.e., number range, counting in 1s, 2s, decade crossings, etc.) and thus difficulty. The starting points align with the UK Curriculum (Department for Education, 2021) reflecting that during the Early Years Foundation Stage, children up to 5 years (Reception year) should learn to verbally count beyond 20 (reflected by starting point 1), and recognise the pattern of the counting system (e.g., even numbers patterns: 2, 4, 6, ...; odd number patterns: 3, 5, 7, ... reflected by starting point 2). Children in Year 1 should learn to count, read, and write numbers up to 100 in numerals, count in multiples of 2s, 5s and 10s, as well as identify and represent numbers using pictorial representations, including the number line (Department for Education, 2021). This is reflected in starting points 2 to 7, in which children have to count in multiples of 2s, 5s, and 10s within the number range of up to 119 involving or not involving decade crossings. This way, the game facilitates children's understanding of the ordinal and spatial relations between numbers by working with number lines while also strengthening the association between Arabic numbers and the corresponding number word (e.g., 10 = "Ten"). Children can learn to count in 2s, 5s, and 10s, while exploring the number range from 1 to 120, including trials involving decade crossings. Additionally, children practise some arithmetic when buying items in the shop although this is not the primary learning opportunity of the game. All these features make the game align well with the aims of the UK curriculum for Early Years Foundation Stage and Year 1.

Table 1: Starting points of the game.

| Starting points | Levels | Numerical Content |
|------------------|--------|---|
| starting point 1 | 1-10 | Counting in 1s between 1 and 20 |
| starting point 2 | 11-17 | Counting in 2s between 2 and 24 |
| starting point 3 | 18-24 | Counting in 10s between 10 and 120 |
| starting point 4 | 25-31 | Counting in 5s between 5 and 60 |
| starting point 5 | 32-38 | Counting in 2s odd numbers between 1 and 23 |
| starting point 6 | 39-78 | Counting in 1s between 21 and 119 not crossing tens |
| starting point 7 | 79-118 | Counting in 1s between 27 and 123 crossing tens |

2.3 Adapting the Difficulty of Levels According to Players' Performance

Irrespective of the starting point chosen, we implemented an adaptive algorithm to determine the player's progression through the game. The algorithm keeps the player in a zone of optimal challenge by presenting more difficult levels in case of a given number of correct responses and easier levels in case of incorrect responses (see Table 2a and Table 2b for details). An incorrect response is when the player puts the number on the wrong carriage. The train will not leave the station unless all the numbers are on the right carriages. There is no time limit for the player to correct their responses. For players that commit two or fewer errors, there is a 10% chance they play rounds on the same level, a 60% chance they go to the very next more difficult level, a 10% chance they go to two levels up and a 5% chance they go three levels up. For players who give two or more incorrect responses, there is a 60% chance they play rounds on the same level, a 25% chance they go to the previous level, a 10% chance they go two levels down and a 5% chance they go three levels down.

Table 2a: The adaptation algorithm in case of ≤ 2 incorrect responses in a round (each level has 5 rounds). The algorithm determines the difficulty of the next round based on the present round player performance

| | | | | | |
|--|-----|-----|-----|-----|-----|
| In case the player had ≤ 2 incorrect responses (i.e., clicking on the locomotive to leave the station, but numbers were not in the correct order) in the current round, the algorithm will direct the player to a specific level for the next round with the following probabilities. | | | | | |
| Next Level | L-1 | L | L+1 | L+2 | L+3 |
| Percentage of next-level selection | 0% | 10% | 60% | 25% | 5% |

Table 2b: The adaptation algorithm, in case of > 2 incorrect responses in a round (each level has 5 rounds. The algorithm decides the difficulty of the next round based on the present round player performance.)

| | | | | | |
|--|-----|-----|-----|-----|-----|
| In case, the player had > 2 incorrect responses in the current round, the algorithm would direct them to a specific level for the next round with the following probabilities. | | | | | |
| Next Level | L-3 | L-2 | L-1 | L | L+1 |
| Percentage of next-level selection | 5% | 10% | 25% | 60% | 0% |

2.4 Game Elements for Motivation

Malone (1980) described three elements that can make a game fun, and increase intrinsic motivation: *challenge*, *fantasy*, and *curiosity* (Malone, 1980). To be motivating, a game needs to present a challenge to the player with goals and uncertain outcomes, that can engage a player's self-esteem. Moreover, fantasy (i.e., a game can stimulate a player's imagination with its story) can make a game more interesting because the game can evoke images of scenarios, situations or objects that can engage players. Lastly, a game is more engaging when it stimulates players' sensory and cognitive curiosity with novel and surprising content. According to Malone's taxonomy, we considered some of these elements when designing the game.

The adaptive algorithm manipulates the difficulty level to keep the player in a zone of optimal challenge. Moreover, the algorithm adds a surprise element as children who correctly responded in a round may suddenly face much more difficult levels. It is important to notice that this surprise of being presented with much more difficult levels does not apply to children who are already struggling with the current level to avoid any frustration. Another element that makes *Number Express* challenging is its *goal*. The *goal* of *Number Express* is to number the carriages correctly to allow the train to leave the station. When players do this correctly, they earn virtual coins to spend in the *shop*, which could be considered an additional goal itself. Children's desire to personalize their character and thus buy items in the shop may further motivate them to play the game for longer. Moreover, buying items in the shop offers more opportunities for learning mathematics, such as comparing the cost of an item to the coins a player has and calculating how many more coins are needed to buy a specific item.

For designing this game, we hired a professional graphic designer to create illustrations that stimulate children's fantasy. We also chose to present shop items that are frequently liked by both boys and girls. The detailed background and the possibility of choosing and naming their character to play the game should also further stimulate children's fantasy and immerse them in the game's narrative.

To stimulate cognitive curiosity, we implemented a *two-stage informative feedback system*. In particular, when the player puts the numbers on carriages in the wrong order, the wrongly ordered number labels move to the top of the screen again, and a visual and vocal message appears to inform the player that the indicated order was not correct (i.e., stage 1, see Figure 5, left chart). At this point, the player can associate the numbers with the respective carriages again. In case, the player gets the order wrong again, a number line will appear under the train so they can use it to order the numbers correctly (i.e., stage 2, see Figure 5, right chart). Stage 2 shall guarantee that all children can place the numbers in the correct order by matching the numbers on the carriages to those on the number line. This feedback provides players, including those struggling with number ordering, with enough information to order the numbers accurately (Figure 5).

Overall, the game's goal and the adaptative algorithm make the game *challenging* for children. The visual aesthetics and the narrative stimulate children's *fantasy* of being the train conductor who needs to make the train leave the station. Finally, the two-step informative feedback system ensures that children get enough information to complete the task at hand.

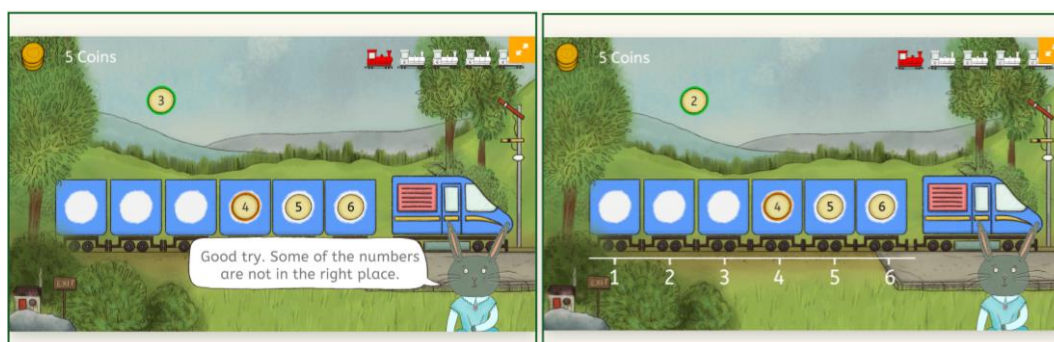


Figure 5: The two-level informative feedback system

3. The Design of the Game

To design *Number Express*, we reviewed the literature on number ordering skills (Devlin et al., 2022; Sasanguie & Vos, 2018; Sella et al., 2020) and on what makes a digital game engaging for children (Malone, 1980; 1981). We considered studies on the importance of the number line as a tool to spatially represent numbers in number games (e.g., Siegler & Ramani, 2008). This evidence guided the design of the game. Accordingly, the numbers on the screen are presented in linear order and equidistantly to each other with numerical magnitude increasing from left to right (i.e., 1-2-3-4...). Moreover, we set the levels of difficulty according to the national math curriculum for the Early Years Foundation Stage and Year 1 in the UK. Furthermore, we had a professional graphic designer create the game scenarios and characters mimicking the style of popular books in nurseries and primary Schools in the UK to make sure the illustrations were familiar to the children and suitable for their age. Finally, the first version of the game was modified according to the received feedback from researchers, teachers, and an education consultant specialising in the learning and teaching of primary mathematics to further improve the game based on their feedback. Then we run a pilot study, which is described in the next section.

4. Initial Incremental Changes of the Game

After preparing the first version of the game, we followed a design-based research method to gather initial feedback. We sought input from academic colleagues with cognitive and pedagogical backgrounds, a primary school teacher, and an educational consultant specializing in Early Years mathematics education. We initially presented the game to a group of colleagues in a mathematics pedagogy seminar at the Department of Mathematics Education, Loughborough University. Their primary suggestion was to implement an option for players to choose the level of difficulty and number range at the start of the game. This feature prevents players from getting bored by playing through the easier levels before reaching more challenging ones.

In the second step, we consulted additional colleagues from the Department of Mathematics Education for further feedback. They recommended introducing larger number intervals, ranging from 20 to 120, without crossing a decade before introducing intervals that do.

In the third step, we shared the game with a primary school teacher. He found the game engaging, noting that the learning goals were clear and aligned with the UK curriculum. His suggestion was to group the game levels

in a way that prevents children, parents, or teachers from feeling confused by too many choices when selecting them.

As the fourth step, we presented the game to an Early Years consultant. She advised us on how to improve the game's graphics, making them less distracting and more meaningful for children. In particular, she suggested we use an image of a stack of coins to mark the points earned instead of a single coin, to be mathematically coherent. Indeed, having a number greater than 1 next to an image of a single coin could confuse children. Additionally, she recommended adjusting some levels to focus more on number ordering (we initially included levels with addition and subtraction). Finally, she suggested a slight modification to how numbers reappear on the screen after a wrong response from the children. Initially, when children placed the numbers in the wrong order, all the numbers in incorrect positions appeared at the top of the screen to be re-ordered by the children. That could be confusing. Therefore, she suggested making the numbers in the wrong positions reappear one at a time. This way, children can focus on positioning one number at a time.

Following this initial phase, we felt the game was ready for piloting it with children.

5. Evaluation of the Game's Playability and User Experience

5.1 Method

5.1.1 Participants

Forty-five preschool children (from 4 to 6 years of age; 26 boys) participated in the study during their last kindergarten year. Children were randomly separated into two groups one playing *Number Express* on a desktop PC (18 children) and one playing the game on a tablet (27 children).

5.1.2 Design

Children played four 15-minute sessions in small groups in a separate room supervised by a trained research assistant. At the beginning of the first session, children watched a video tutorial to understand how to play the game. During the other three sessions, they watched the video tutorial or had some help remembering how to play the game, but only when they asked for it. Trained research assistants supported the children in playing, filled in the observation checklist (see Table 3) during each play session, and, at the end of the last play session, asked children a few questions about their user experience (see Table 4). Research assistants kept their interaction with children to a minimum, resembling future user contexts in which teachers would set up play sessions for multiple children without the necessity of constantly supervising them.

5.1.3 Materials

We aimed to evaluate the game's playability and user experience as "a quality attribute that assesses how easy user interfaces are to use" (Nielsen, 2012, paragraph 1) across five components (Nielsen, 2012): i) *learnability* (e.g., how easy it is for users to accomplish basic tasks the first time they encounter the game), ii) *memorability* (e.g., when users return to the design after a period of not using it, how easily they re-establish proficiency), iii) *errors* (e.g., how many errors users make, how severe these errors are, and how easily they recover from the errors), iv) *efficiency* (e.g., once users have learned the game, how quickly can they perform tasks), and v) *satisfaction* (e.g., how pleasant it is to use the game). In this vein, we developed an observation checklist and an interview-based questionnaire, which were adapted from Ismail and colleagues (2011).

The observation checklist assessed *Learnability*, *Memorability*, *Effectiveness*, *Errors*, *Self-corrections*, and *Efficiency* (Table 3). The *Learnability* item captures the number of times children requested to watch the video tutorial. During the first session, all children watched the video tutorial once. From the second session onwards, the research assistants recorded the number of times a child asked to watch the video tutorial.

In the *Memorability* item, children were awarded a score between 1 and 10 based on how well they remembered how to play the game (1= "the child did not remember how to play the game"; 10= "the child remembered well how to play the game with no need to watch the video tutorial").

The *Effectiveness* subscale entailed eight items to investigate whether children were confident in using some technical features of the game (e.g., how to start playing, and how to choose the train conductor). All items were scored from 1 to 10 based on how well children used the game's features (see Table 3) where 1 indicates poor use and 10 indicates excellent use. We removed Item 3 ("The child did not need help to write the train conductor's name.") from the analysis because some tablets failed to display the keyboard to type the name of

the conductor, and some children just skipped this step and went directly to play the game. For each child in each session, we calculated the average across the seven items as an overall effectiveness index.

The *Error* item assessed how frequently children committed errors during the play sessions, where 10 was the best score (“the child did not commit any error while playing the game”) and 1 was the worst score (“the child committed many errors while playing the game”).

In the *Self-correction* item, children were awarded a score between 1 and 10 based on how confident they were in correcting their errors, where 10 was the best score (“The child is confident in correcting their errors”) and 1 was the worst score (“The child did not know how to correct their errors”).

The *Efficacy* item assessed whether the number line, which appeared in the case of two consecutive errors in numbering carriages, was an effective scaffolding tool to prompt children to order numbers correctly. No scores were assigned in case the number line did not appear on the screen, meaning the children did not commit two consecutive errors. A score of 10 was the maximum score (“The child was able to put the numbers in order using the number line”) and 1 was the minimum score (“The child was not able to put the numbers in order even using the number line”).

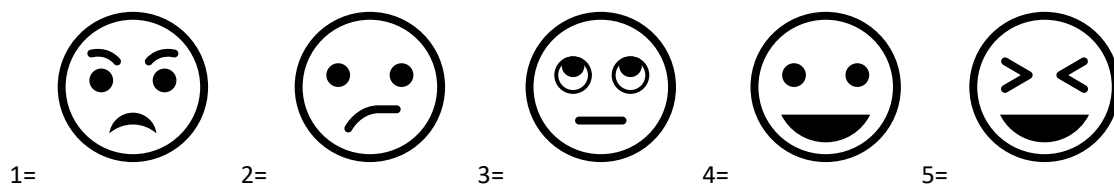
Table 3: The Usability Observation Checklist

| Component | Item |
|-----------------|--|
| Learnability | The number of times a child requested to see the video tutorial. |
| Memorability | The child remembered how to play without asking to watch the video tutorial or seeking support from the adult. |
| Effectiveness | The child understood how to start the game. |
| | The child could select the train conductor. |
| | The child did not need help to write the train conductor’s name. |
| | The child knew how to move the number tag through the screen. |
| | The child understood how to make the train leave. |
| | The child understood how to enter the shop. |
| | The child could buy items in the shop. |
| | The child could come back to the game after having been in the shop. |
| Errors | Errors. |
| Self-correction | The child knows how to correct themselves. |
| Efficacy | The number line helps the child to put the number in the right order. |

Finally, we evaluated the children’s *User experience* using an interview-based questionnaire at the end of the fourth play session. Children had to point to emoji faces corresponding to values from 1 to 5 to assess their response agreement with 5 statements (see Table 4).

Table 4: The Interview-based Questionnaire

| Statements |
|--|
| 1. The game was fun. |
| 2. The game was easy to play. |
| 3. I like the game characters. |
| 4. I want to play this game again. |
| 5. I would like to play this game at home. |



5.1.4 Data analysis

Data analysis was performed using the statistical software JASP (JASP Team, 2023), and plots were created using R (R Core Team, 2023). We mainly analysed changes in scores between the two groups [PC, Tablet] across the four play sessions by running mixed ANOVAs. We applied the Greenhouse-Geisser correction for p-values when the sphericity assumption was violated.

6. Results

6.1 Learnability

The *Learnability* score indicated the number of times children requested to watch the video instructions before playing the game. We excluded the data from the first session because the research assistant showed the video instructions to all children once. Although possible, none of the children requested to view the video again during the first play session. Note that none of the children requested to view the video more than once in a session, constraining all values between 0 and 1, and that, in the fourth session, no children requested to watch the video instructions before playing.

We analysed the *learnability* score in a mixed ANOVA with Session [2, 3] as within-participants factor and Device [PC, Tablet] as between-participants factor. The main effects of Session [$F(1, 43)=12.39, p<.001, \eta^2_p=0.224$] and Device [$F(1, 43)=7.78, p=.008, \eta^2_p=0.153$] were significant. In contrast, the interaction between Session and Group [$F(1, 43)=1.56, p=.218, \eta^2_p=0.035$] was not. Overall, children in the tablet group requested to watch the video more often ($M=0.36, SD=0.29$) than children in the PC group ($M=0.13, SD=0.26$). Children displayed lower scores from the second to the third session, as illustrated in Figure 6. To further explore the effect of Session, we also run two t-tests with Bonferroni correction (2 comparisons) comparing the scores in the second session and the third session against 0 (i.e., the fourth session), separately. The scores in the second session differed from 0 ($M=0.36, SD=0.48; t(45)=4.93, p<.001$) whereas the scores in the third session did not ($M=0.09, SD=0.29; t(45)=2.07, p=.088$).

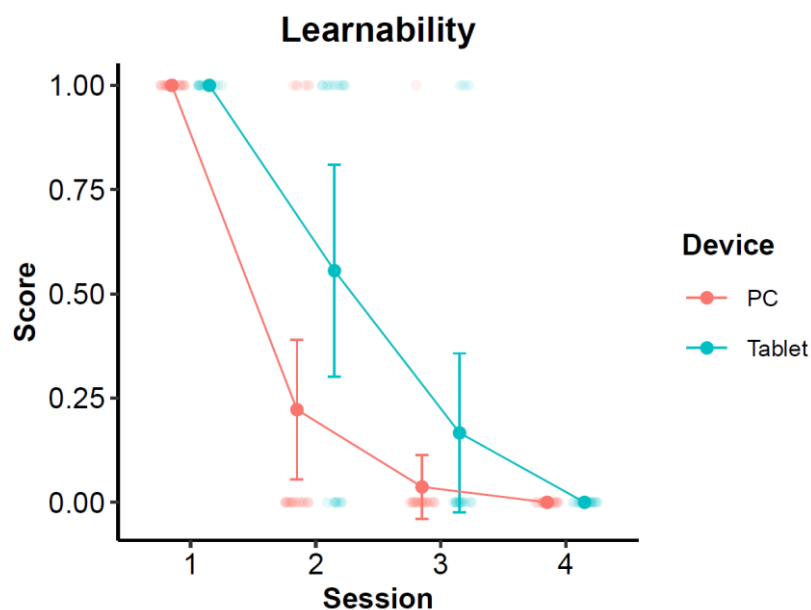


Figure 6: Means of the number of times (y-axis) children in the two groups (PC and Tablet) requested to watch the video instructions before playing the game across the four sessions (x-axis). The transparent points represent individual scores and error bars represent 95% CI

6.2 Memorability

The *memorability* score reflects children's ability to remember how to play the game after the first session, with higher values indicating a better ability to engage with the game without additional support from assistants. Please note that all children received a score of 10 in the fourth playing session.

We analysed the *memorability* score using a mixed ANOVA with Session [2, 3] as within-participants factor and Device [PC, Tablet] as between-participants factor. The main effects of Session [$F(1, 43)=15.14, p<.001$,

$\eta^2_p=0.260$] and Device [$F(1, 43)=8.04, p=.007, \eta^2_p=0.153$] were significant. However, the interaction between Session and Group [$F(1, 43)=1.16, p=.297, \eta^2_p=0.025$] was not significant. Overall, children in the tablet group showed lower memorability scores ($M=6.61, SD=2.4$) than children in the PC group ($M=8.65, SD=2.33$). To further explore the effect of Session, we also run two t-tests with Bonferroni correction (2 comparisons) comparing the scores in the second session and the third session against 10 (i.e., the fourth session), separately. The scores in the second session differed from 10 ($M=6.51, SD=4.24; t(45)=5.52, p<.001$) whereas the scores in the third session did not ($M=9.16, SD=2.48; t(45)=2.29, p=.054$). The pattern of results indicates that children progressively remember how to play the game, highlighting a more pronounced improvement from sessions 2 to 3, which reached a plateau in the third session, as illustrated in Figure 7.

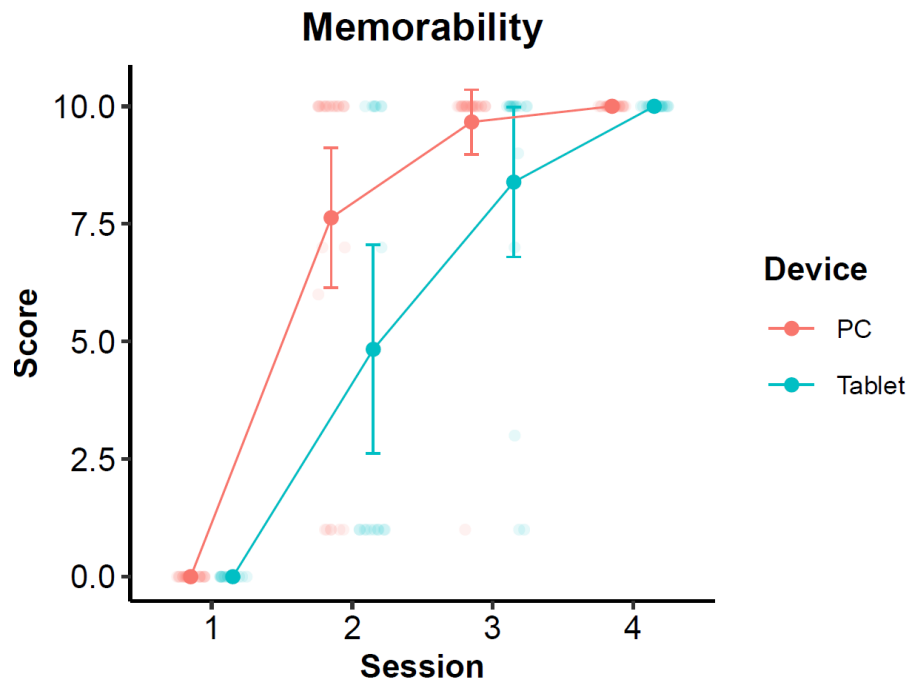


Figure 7: Mean scores of Memorability (y-axis) in the two groups (PC and Tablet) across the sessions (x-axis). The transparent points represent individual scores and error bars represent 95% CI

6.3 Effectiveness

We analysed the mean of the *Effectiveness* subscale in a mixed ANOVA with Session [1, 2, 3, 4] as within-participants factor and Device [PC, Tablet] as between-participants factor. The main effect of Session [$F(1.34, 57.8)=155.96, p_{[gg]}<.001, \eta^2_p=0.784$], the main effect of Group [$F(1, 43)=5.8, p=.02, \eta^2_p=0.119$], and the interaction between Session and Group [$F(1.34, 57.8)=5.06, p_{[gg]}=.019$] were significant. To further explore the significant interaction, three ANOVAs with Device as between-subjects factor and Session as within-subjects factors separately for session 1 and 2, 2 and 3, 3 and 4. The interactions between Device and Session were significant for the comparison of Session 1 and 2 ($p=0.019$), and 3 and 4 ($p=0.022$), whereas the tablet group seemed to show larger improvement in their memorability scores passing from session 1 to 2 and from session 3 to 4 compared to the PC group. We further investigated the main effect of Session by running post-hoc pairwise t-tests with Bonferroni correction (3 comparisons) comparing session 1 vs 2 ($t(44)=11.84, p<.001$), 2 vs 3 ($t(44)=3.67, p=.002$), and 3 vs 4 ($t(44)=2.37, p=.075$). Children became more effective in interacting with the game, with a more relevant improvement from sessions 1 to 2, which reached a plateau in the third session, as illustrated in Figure 8.

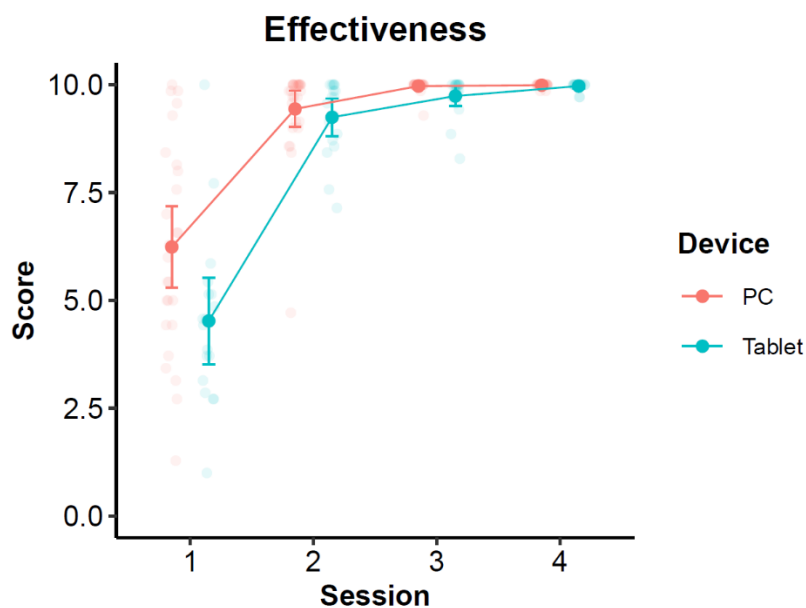


Figure 8: Mean scores of Effectiveness (y-axis) in the two groups (PC and Tablet) across the four sessions (x-axis). The transparent points represent individual values, and the error bars represent 95%CI

6.4 Error

We analysed the error item in a mixed ANOVA with Session [1, 2, 3, 4] as within-participants factor and Device [PC, Tablet] as between-participants factor. The main effect of Session ($F(1.9, 81.5)=89.27, p_{[gg]}<.001, \eta^2_p=0.675$) and the main effect of Group ($F(1, 43)=4.85, p=.033, \eta^2_p=0.101$) were both significant. In contrast, the interaction between Session and Group ($F(1.9, 81.5)=1.16, p_{[gg]}=.318, \eta^2_p=0.026$) did not reach statistical significance. Overall, children playing on the tablet displayed higher scores on the item Error ($M=2.89, SD=1.34$) than children playing on the PC ($M=2.19, SD=0.78$). To further explore the main effect of Session, we ran post-hoc pairwise t-tests with Bonferroni correction (4 comparisons) comparing session 1 vs 2 ($t(44)=8.54, p<.001$), 2 vs 3 ($t(44)=4.02, p=.001$), 3 vs 4 ($t(44)=2.07, p=.176$), 2 vs 4 ($t(44)=4.32, p<.001$). The pattern of results indicates that children committed fewer errors with a clear reduction from session 1 to 2 and reached a plateau in the third session, as illustrated in Figure 9.

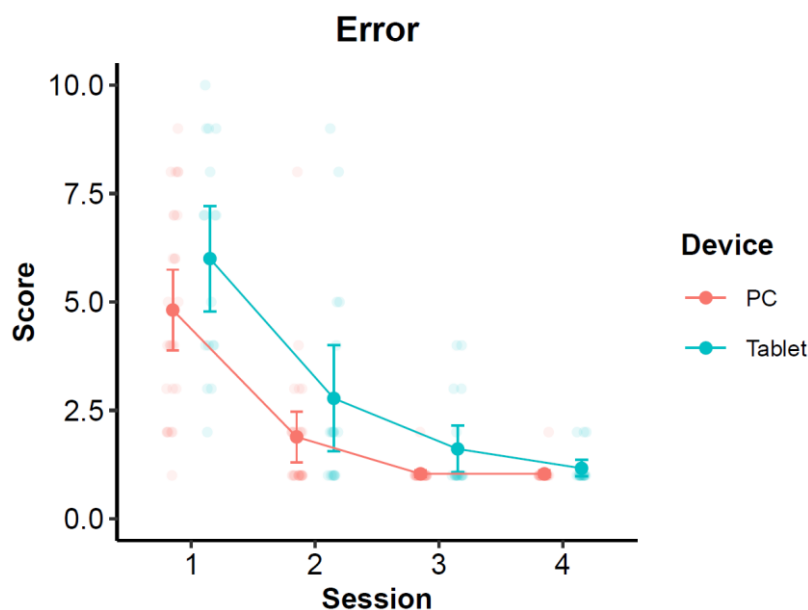


Figure 9: Means scores of Error item (y-axis) in the two groups (PC and Tablet) across the four sessions (x-axis). The transparent points represent individual values, and the error bars represent 95% CI

6.5 Self-Correction and Efficiency

The self-correction and the efficiency items entailed very low variability and reduced sample size. Therefore, we only reported descriptive statistics for these measures. Most children easily self-correct themselves, as highlighted by the high scores in the self-correction item across the four play sessions (Session 1: $M=9.56$, $SD=1.67$; Session 2: $M=9.67$, $SD=1.38$; Session 3: $M=10$, $SD=0$; Session 4: $M=10$, $SD=0$). Similarly, fewer children committed two consecutive errors in ordering numbers and were shown the number line as a scaffolding tool to order numbers correctly. Children were very effective in using the number line feedback as indexed by high scores in the Efficiency item (Session 1: $M=8.86$, $SD=2.66$, $n=22$; Session 2: $M=8.88$, $SD=2.94$, $n=16$; Session 3: $M=9.29$, $SD=1.9$, $n=14$; Session 4: $M=9.92$, $SD=0.29$, $n=12$).

6.6 User Experience

The scores for the five questions were very high and close to the maximum score of 5 (Question 1: $M=4.76$, $SD=0.71$; Question 2: $M=4.56$, $SD=0.66$; Question 3: $M=4.36$, $SD=0.96$; Question 4: $M=4.49$, $SD=1.04$; Question 5: $M=4.69$, $SD=0.67$), meaning that children enjoyed playing the game.

7. Conclusion and Future Perspectives

Number Express is a digital game designed to support basic numerical skills in pre- and primary-school children. We designed the game to accommodate crucial aspects of basic numerical development such as ordinality, the spatial representation of numbers (i.e., number line), and counting in 1s, 2s, 5s, and 10s.

In the current study, we describe how we sought initial feedback on the game following a design-based research method and a study piloting the game with a group of four to six-year-old children to evaluate its usability and likability.

The final version of the game was developed implementing an iterative design-based research approach where we consulted with expert educational practitioners and incorporated their feedback over several development circles which led to significant changes and adaptations of the original (design) ideas for the game.

During the pilot study, we investigated the usability of the game. Overall, the game proved to be simple to play, with instructions that were easy to memorise, the number line tool was effective in supporting the number ordering activity and the children enjoyed playing the game.

We found that while the children involved in the study watched a video tutorial during the first session of the study, only half of the children needed to watch the video tutorial during the second session of the study to remember how to play the game. By the third session, almost all the children remember how to play the game without the need to watch the video tutorial.

Since the first session, children have shown a good understanding and use of the elements of the game that improved over the sessions, as the analysis of the Effectiveness subscale showed. Children became more effective in interacting with the game, with a more relevant improvement from sessions 1 to 2, which reached a plateau in the third session. By the second session, children used all the features of the game optimally.

Less than half of the children needed the help of the number line to put the number in the right order. For them, this tool proved to be very effective.

Lastly, the results of the user experience questionnaire indicated a high level of satisfaction among the participants.

Of course, future research on the effectiveness of the game is needed. In the next step of the development process, we aim to evaluate the game's educational effectiveness in a controlled large-scale intervention study from which we expect a facilitating effect of playing the game on children's mastery of basic numerical skills. We consider the further development of the game an ongoing process during which we may keep adapting and optimising game elements to maximise its user experience as well as educational effectiveness according to incoming feedback and empirical evidence.

Taken together, the findings from this study indicate that Number Express is straightforward to engage with, features easily memorable instructions, and incorporates an effective number line tool that supports number ordering skills.

Acknowledgements

We want to thank all colleagues who gave us feedback on the game: Kinga Morsanyi, Tim Jay, Dave Hewitt, Mathematical Cognition Group (Loughborough University), Mathematical Pedagogy Group (Loughborough University), Martin Tillbrook, and Helen Williams. We also want to thank the children who played our game during the pilot study.

Ethic statement: the project has been approved by the Ethical Committee for the Psychological Research of the University of Padova (5403).

Funding: this work was (partially) supported by UKRI Economic and Social Research Council [grant number ES/W002914/1].

Data availability statement: data will be made available upon request.

AI statement: we have not used an AI tool.

References

- Annetta, L. A., Minogue, J., Holmes, S. Y., and Cheng, M. T. 2009. Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers and Education*, 53(1), 74–85. Available at: <https://doi.org/10.1016/j.compedu.2008.12.020>
- Atkinson, A. L., Hill, L. J. B., Pettinger, K. J., Wright, J., Hart, A. R., Dickerson, J., and Mon-Williams, M. 2022. Can holistic school readiness evaluations predict academic achievement and special educational needs status? Evidence from the Early Years Foundation Stage Profile. *Learning and Instruction*, 77, 101537. Available at: <https://doi.org/10.1016/J.LEARNINSTRUC.2021.101537>
- Balladares, J., Miranda, M. and Cordova, K., 2024. The effects of board games on math skills in children attending prekindergarten and kindergarten: A systematic review. *Early Years*, 44(3-4), pp.710-734.
- Benavides-Varela, S., Callegher, C.Z., Fagiolini, B., Leo, I., Altoè, G. and Lucangeli, D. 2020. Effectiveness of digital-based interventions for children with mathematical learning difficulties: A meta-analysis, *Computers & Education*, 157, p.103953.
- BITESIZE, 2024. Maths EYFS - Bud's Number Garden. [online] Available at: <https://www.bbc.co.uk/bitesize/articles/zd4b382> [Accessed 30 September 2024]
- Crittenden, W. F., Biel, I. K., and Lovely III, W. A. 2019. Embracing digitalization: Student learning and new technologies. *Journal of marketing education*, 41(1), 5-14.
- Davis-Kean, P. E., Domina, T., Kuhfeld, M., Ellis, A., and Gershoff, E. T. 2022. It matters how you start: Early numeracy mastery predicts high school math course-taking and college attendance. *Infant and Child Development*, 31(2). Available at: <https://doi.org/10.1002/icd.2281>
- Department for Education. 2021. National curriculum in England: mathematics programmes of study. Available at: <https://www.gov.uk/government/publications/national-curriculum-in-england-mathematics-programmes-of-study/national-curriculum-in-england-mathematics-programmes-of-study>. Accessed: April 2023.
- Department for Education. 2021. Statutory framework for the early years foundation stage. Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/974907/EYFS_framework_-_March_2021.pdf (Accessed: April 2023)
- Devlin, D., Moeller, K., Reynvoet, B., and Sella, F. 2022. A critical review of number order judgements and arithmetic: What do order verification tasks actually measure? *Cognitive Development* (Vol. 64). Elsevier Ltd. Available at: <https://doi.org/10.1016/j.cogdev.2022.101262>
- Devlin, D., Moeller, K., Xenidou-Dervou, I., Reynvoet, B. and Sella, F., 2024. Familiar Sequences Are Processed Faster Than Unfamiliar Sequences, Even When They Do Not Match the Count-List. *Cognitive Science*, 48(7), p.e13481.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., Japel, C., Cordray, D., Ginsburg, H., Grissmer, D., Lipsey, M., Raver, C., Sameroff, A., and Zill, N. 2006. School Readiness and Later Achievement, *Developmental psychology*, 43(6), 1428.
- Duncan, R. J., Duncan, G. J., Stanley, L., Aguilar, E., and Halfon, N. 2020. The kindergarten Early Development Instrument predicts third grade academic proficiency, *Early Childhood Research Quarterly*, 53, 287–300. Available at: <https://doi.org/10.1016/J.ECRESQ.2020.05.009>
- Education.com, 2024. Jumpy: Counting by Ones From 1 to 100. [online] Available at: <https://www.education.com/game/jumpy-counting-by-ones-from-1-to-100/> [Accessed 30 September 2024]
- Gelman and Gallistel. 1978. *The Child's Understanding of Number*. Cambridge, MA. Harvard University Press.
- INSERM-CEA Cognitive Neuroimaging Unit. 2004. The number race. Available at: <http://www.thenumberrace.com/nr/home.php>.
- INSERM-CEA Cognitive Neuroimaging Unit. 2011. Number catcher. Available at: <http://www.thenumbercatcher.com/nc/home.php>.

- Ismail, M., Diah, N. M., Ahmad, S., Kamal, N. A. M., and Dahari, M. K. M. 2011. Measuring usability of educational computer games based on the user success rate. *SHUSER 2011 - 2011 International Symposium on Humanities*, in Science and Engineering Research, 56–60. <https://doi.org/10.1109/SHUSER.2011.6008500>
- JASP Team. 2023. JASP (Version 0.16) [Computer software]. Retrieved from <https://jasp-stats.org/>
- Jordan, N. C., Hanich, L. B., and Kaplan, D. 2003. A Longitudinal Study of Mathematical Competencies in Children With Specific Mathematics Difficulties Versus Children With Comorbid Mathematics and Reading Difficulties. *Child Development*, 74(3), 834–850. Available at: <https://doi.org/10.1111/1467-8624.00571>
- Käser, T., Baschera, G.M., Kohn, J., Kucian, K., Richtmann, V., Grond, U., Gross, M. and von Aster, M., 2013. Design and evaluation of the computer-based training program *Calcularis* for enhancing numerical cognition. *Frontiers in psychology*, 4, p.489.
- Ke, F. 2009. A qualitative meta-analysis of computer games as learning tools. *Handbook of research on effective electronic gaming in education*, 1-32.
- Klawe, M. 1999, June. Computer games, education and interfaces: The E-GEMS project. In *Graphics interface* (pp. 36-39).
- Malone, T. W. 1980, September. "What makes things fun to learn? Heuristics for designing instructional computer games in Proceedings of the 3rd ACM SIGSMALL symposium and the first SIGPC symposium on Small systems (pp. 162-169).
- Morsanyi, K., O'Mahony, E. and McCormack, T., 2017. Number comparison and number ordering as predictors of arithmetic performance in adults: Exploring the link between the two skills, and investigating the question of domain-specificity. *Quarterly Journal of Experimental Psychology*, 70(12), pp.2497-2517.
- Nielsen, J. 2012, January, 3. *Usability: Introduction to Usability*. Available at: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/> (Accessed: April 2023).
- Ninaus, M., Kiili, K., McMullen, J. and Moeller, K., 2017. Assessing fraction knowledge by a digital game. *Computers in Human Behavior*, 70, pp.197-206.
- Pareto, L., Arvemo, T., Dahl, Y., Haake, M., & Gulz, A. 2011. A teachable-agent arithmetic game's effects on mathematics understanding, attitude and self-efficacy in *Lecture Notes in Computer Science* (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 6738 LNAI, 247–255. Available at: https://doi.org/10.1007/978-3-642-21869-9_33
- R Core Team. 2023. R: A language and environment for statistical computing [Computer software]. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Sarnecka, B. W., Goldman, M. C., and Slusser, E. B. 2015. *How counting leads to children's first representations of exact, large numbers*. The Oxford handbook of numerical cognition, 291-309.
- Sasanguie, D., and Vos, H. 2018. About why there is a shift from cardinal to ordinal processing in the association with arithmetic between first and second grade. *Developmental Science*, 21(5), e12653.
- Sella, F., Berteletti, I., Lucangeli, D., and Zorzi, M. 2017. Preschool children use space, rather than counting, to infer the numerical magnitude of digits: Evidence for a spatial mapping principle. *Cognition*, 158, 56-67.
- Sella, F., Onnivello, S., Lunardon, M., Lanfranchi, S., and Zorzi, M. 2021. Training basic numerical skills in children with Down syndrome using the computerized game "The Number Race". *Scientific Reports*, 11(1), 2087.
- Sella, F., Tressoldi, P., Lucangeli, D., and Zorzi, M. 2016. Training numerical skills with the adaptive videogame "The Number Race": A randomized controlled trial on preschoolers. *Trends in Neuroscience and Education*, 5(1), 20-29.
- Siegler, R. S., and Ramani, G. B. 2008. Playing linear numerical board games promotes low-income children's numerical development. *Developmental science*, 11(5), 655-661.
- Watts, T. W., Duncan, G. J., Siegler, R. S., and Davis-Kean, P. E. 2014. The Groove of Growth: How Early Gains in Math Ability Influence Adolescent Achievement. *Society for Research on Educational Effectiveness*. SREE Spring 2014 Conference Abstract Template.
- Wilson, A.J., Dehaene, S., Pinel, P., Revkin, S.K., Cohen, L. and Cohen, D., 2006. Principles underlying the design of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioral and brain functions*, 2, pp.1-14.