International Journal of Child-Computer Interaction 27 (2021) 100222

Contents lists available at ScienceDirect

International Journal of Child-Computer Interaction

journal homepage: www.elsevier.com/locate/ijcci

Leveraging tangible interfaces in primary school math: Pilot testing of the Owlet math program[☆]

Lauren Zito [a], Jennifer L. Cross [b], Bambi Brewer [c], Samantha Speer [a], Michael Tasota [a], Emily Hamner [a,*], Molly Johnson [a], Tom Lauwers [c], Illah Nourbakhsh [a]

[a] CREATE Lab, Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, United States of America
[b] Center for Engineering Education and Outreach, Tufts University, 200 Boston Avenue, Suite G810, Medford, MA 02155, United States of America
[c] BirdBrain Technologies LLC, 544 Miltenberger St, Pittsburgh, PA 15219, United States of America

ARTICLE INFO

Article history:
Received 1 September 2020
Received in revised form 10 November 2020
Accepted 13 November 2020
Available online 20 November 2020

Keywords:
Digital manipulative
Math manipulative
Primary math education
Tangible user interface

ABSTRACT

This paper presents pilot testing of Owlet, a math program based on two original, tangible interface devices for primary school math and their accompanying apps. We built on prior work that demonstrated promising outcomes regarding manipulatives in math education and tangible user interfaces in a variety of applications. The Owlet program was pilot tested in ten classrooms, spanning students ages 5 to 11. We found that teachers used the exploratory activities to introduce the program, and other activities to encourage differentiated, student-paced practice of math concepts. Students were interested and engaged in using Owlet during the pilot tests, leading to student driven generation of challenges. Through development, Owlet, as a whole program, spanned more math concepts by prioritizing flexibility in one tangible interface and concreteness in the other. Our findings highlight the strengths and weaknesses of each tangible device.

1. Introduction

While STEM education is seen as a critical driver of economic growth and innovation, many US students lack critical mathematics skills [1]. On the 2015 National Assessment of Education Progress (NAEP), 75% of 12th grade students (17 to 18 years old) performed below the proficient level [2]. Younger students performed only slightly better; approximately 67% of 8th graders (13 to 14 years old) and 60% of 4th graders (9 to 10 years old) are not proficient in math [3]. It is clear from these surveys that K-12 mathematics education is failing the majority of US students.

Due to the importance of mathematics skills, many studies have investigated the application of digital technology and software to support mathematics education [4]. Education software has been developed for a variety of applications, which can broadly be categorized as "drill-and-practice", interactive/educational games, and exploratory tools [5]. In their review of 46 studies, Li and Ma found small positive effects of computer technology on student mathematics achievement [6].

The use of manipulatives, "objects designed to represent explicitly and concretely mathematical ideas that are abstract", has also been shown to aid in student math understanding [7]. However, the use of manipulatives is mediated by teacher beliefs of their utility and frequently requires additional effort from the students in order to extract mathematical ideas and meaning from the physical tools [7,8]. We address these issues through the creation of tangible interfaces that expand the range of concepts explorable by students.

Since 2016, our research group has been working on the development of an elementary math program utilizing tangible user interfaces, now referred to as Owlet [9]. The Owlet Program combines two distinct tangible interfaces with four tablet apps supported by informational cards depicting usage, troubleshooting, and sample activities. In this paper we present our findings through thematic analysis from one year of pilot testing of the Owlet Program in ten elementary classrooms ranging from grades K-5. Through these pilot tests we aimed to improve the Owlet design using design-based research to address the following research questions:

RQ1. What features of the Owlet program and interfaces are used by teachers in their math instruction?

☆ This material is based upon work supported by the National Science Foundation, United States of America under Grant No. 1831177. This work is part of a Small Business Technology Transfer (STTR) project conducted as a collaboration between Carnegie Mellon University and BirdBrain Technologies LLC.

* Corresponding author.
E-mail addresses: jennifer.cross@tufts.edu (J.L. Cross), bambi@birdbraintechnologies.com (B. Brewer), snspeer@andrew.cmu.edu (S. Speer), ehamner@cmu.edu (E. Hamner), tom@birdbraintechnologies.com (T. Lauwers), illah@cs.cmu.edu (I. Nourbakhsh).

https://doi.org/10.1016/j.ijcci.2020.100222
2212-8689/

RQ2. What processes and activities in the Owlet program might enhance and enable student learning?

RQ3. How do different interface designs affect students' ability to make connections between physical manipulatives, on-screen representations, and math concepts?

All participating teachers constructed their own lessons for the in-school pilot tests based on the Owlet tools, choosing how Owlet would be integrated into their classroom. During the pilots, the research team observed both the instructional practices of the teacher as well as the students' use of the tools.

2. Theoretical framework

There has been a long history of research centered around understanding how students learn symbolic mathematics. One theory that has emerged from this research is concreteness fading [10]. Concreteness fading is a foundational theory that is readily used in math instruction and encourages the presentation of concrete ideas with physical interactions, becoming increasingly abstract, and ending in the presentation of the symbolic representation [10,11]. Many other math instructional theories, like embodied mathematics, use concreteness fading as a key design feature [12].

Concreteness fading consists of the sequential transition between two main representations: concrete and abstract. Often there is a third, transitional representation in concreteness fading, that has been shown to be effective in helping students learn [13,14]. During the first phase, the learner is presented with a concrete, physical representation of the math concept in question. This theory provides a foundation for traditional math manipulatives (Section 2.1) [15]. After being presented with a concrete representation, the learner must undergo a transitional phase to the abstract. This phase of concreteness fading is often supported by the teacher.

Many techniques have been explored for supporting teachers in the transitional phase of concreteness fading. Math software is one such supporting tool, often representing math concepts with pictures and symbols on a computer screen [8]. While this strategy covers the pictorial and symbolic stages of concreteness fading, it leaves out the concrete representation. Owlet is designed to fill in this gap through the use of tangible user interfaces (TUIs), another type of supporting tool designed to include the concrete (Section 2.2). One benefit of TUIs is that they can present multiple external representations simultaneously [16]. With dynamic linking, learners may act on one representation and see the results of those actions in another, which can reduce cognitive load [17].

Owlet draws on related works from math manipulatives, tangible interface design, and embodied child–computer interaction to develop a system that bridges concrete and abstract representations of math concepts by utilizing software, manipulatives, and multiple tangible interfaces.

2.1. Math manipulatives

For several decades, the National Council of Teachers of Mathematics has advocated for the use of manipulatives in math classrooms [18]. Manipulatives provide a way for students to show their thinking, which supports productive mathematical discussions with their teachers and peers [19]. Manipulatives provide a concrete, tangible way to explore math problems and construct mathematical understanding [20]. Students can then build on these concrete experiences and move toward understanding mathematical operations at a more abstract level [21]. More recent research highlights the importance of using dual representations, stating that by combining concrete and abstract representations of mathematical concepts, students are better able to transfer and apply skills in the future [22].

A math manipulative is a physical object or set of objects that can be used to represent a math problem in a concrete way. Examples of manipulatives common in primary school math include counters and base ten blocks. Students can use two-color counters to represent simple addition problems, e.g. two red counters and three yellow counters make five counters in all. Base ten blocks are used to introduce the concept of place value. Students combine ten small cubes to make a "rod" of ten. Similarly, ten rods make a "flat" of 100 small cubes.

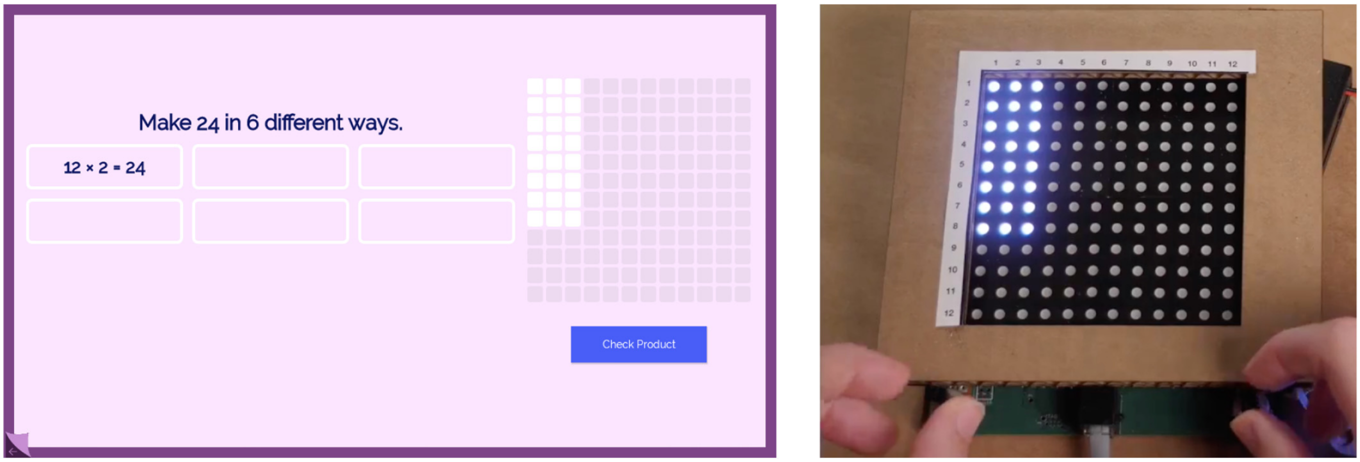
In developing our digital manipulatives, we took inspiration from math manipulatives created by Maria Montessori. Montessori created a wide range of physical manipulatives designed to be used throughout childhood [23]. One example are the Montessori number rods, a series of rods designed to teach students to relate numbers to magnitude. Students sort the rods from smallest to largest and label them with the numerals 1–10. One interesting feature of this manipulative, and other Montessori manipulatives, is that it is self-correcting. When a student places the rods in the wrong order, their error is immediately visually recognizable. The student can then correct the mistake, reinforcing understanding of the underlying math concept. The same manipulative is used over a long period of time in a Montessori classroom, and the design of each manipulative focuses explicitly on a math concept while eliminating details that distract from that concept [24]. We have adopted design goals from Montessori's work to create digital manipulatives that primary school students can use to construct meaning about mathematics.

Math manipulatives provide students with concrete representations of math concepts to aid in understanding and transfer. Owlet not only seeks to provide students with these representations but to link the two representations together through a single interaction, ensuring there are no incongruencies between the concrete and abstract representations. Building upon this, Owlet is designed to not only provide opportunities for students to self correct when they make mistakes with the concrete representations, but also to allow students to practice problems and receive immediate feedback for self correction.

2.2. Tangible user interfaces

Moyer [7] observed that teachers draw a line between exploratory "fun-math" with manipulatives and skill-building "real math" with traditional activities. This mismatch between beliefs about manipulatives and the designed purpose can result in students not building knowledge as intended. She found that students must reflect on their actions with manipulatives to build meaning and that teachers typically need a strong mathematics foundation to make clear connections between the manipulative and the abstract math concept that it is modeling. Additionally, teacher guidance is needed to help students use manipulatives to develop understanding and efficient mathematical strategies without relying long-term on the manipulatives [25]. This echoes the warnings of Ball [26] and Willingham [27] that students cannot learn from manipulatives alone.

As we address these pitfalls through the development of our tangible interface manipulatives, we draw inspiration from a growing movement of tangible user interface exploration. These tangible user interfaces "give physical form to digital information, employing physical artifacts both as *representations* and *controls* for computational media". The *representation* is the physical display or manifestation of information and the *control* enables the physical manipulation of information in the system [28]. The Owlet manipulatives both control and represent digital information on mathematical concepts, with the controls and representations being both physical and digital.



(a) (b)

Fig. 1. GlowGrid supports arithmetic operations by displaying visual representations on the GlowBoard. As students turn the dials on the GlowBoard (1b), they create equations in the GlowGrid app (1a). Above is an example of multiplication where students turn the left dial to set the number of rows (8) and the right dial to select the number of columns (3), which in turn creates the equation $8 \times 3 = 24$.

Tangible user interfaces have been used in a variety of educational and commercial areas. One commercially available example of a tangible user interface is the OSMO system, which supports activities with traditional math manipulatives [29]. OSMO utilizes a digital-augmented adaptation of play money manipulatives, numeral tiles and counters, and tangrams [30,31]. Schroth et al. found that tangible interfaces also promoted student collaboration and engagement. Scarlatos evaluated a number of educational games which used math manipulatives as tangible user interfaces [32]. Both OSMO and work evaluated by Scarlatos use computer vision systems to track the state of the tangible interface. By contrast, the Owlet manipulatives use non-visual sensors to sense the state of the manipulative and provide feedback to students, and focus on different areas of the Elementary math curriculum.

2.3. Child–computer interaction

Child–computer interaction is a newly emerging field that has recently been centered around research involving theories of embodied cognition and of dynamic child development [33]. The field of child–computer interaction has been trending toward topics relating to tangible devices and education, as well as design, coding, and making [34]. Within these fields, research has been conducted into how tangible user interfaces can be utilized to facilitate better spatial learning [35].

Much of the research described by Antle [33] involves theories that children can use the world around them to off-load their cognition leading to further learning. Off-loading cognition can be done through physical actions, acting alone or on other objects, called complementary actions, for example, improving task performance of multi-digit addition by using pencil and paper [33]. Novack et al. found that it is not only the action but the kind of action that matters for student learning. Action itself produced a shallow understanding of a concept, while gesture created a deeper, more generalizable understanding of math concepts. Furthermore, abstract gestures helped students' ability to generalize better than concrete gestures [36].

There have been many kits involving both physical and digital components that focus on the teaching and fostering of computational thinking. Some of these kits even include tangential focus on math concepts such as basic principles of arithmetic, geometry, and measurement [37]. Many studies have been conducted into the efficacy and design of tangible devices for use in classrooms to engage students in storytelling. One study found that students were more engaged and spent longer using the device when it was used collaboratively between pairs of students [38]. Another study found that the use of a tangible device helped students algorithmically find solutions to a complex problem involving story telling [39]. Other studies have focused efforts on math education facilitated through the use of Scratch programming. Brenton et al. [40] found that students aged 9–10 were not only able to create programs using advanced programming concepts, but were also able to engage in creating programs relating to place value utilizing these concepts.

With Owlet, we aim to add to this growing field of research by exploring the use of tangible devices supported by digital apps in elementary math classrooms. We strove to design Owlet to include not only physical control and action, but meaningful gestures that would create deeper understandings as supported by the findings of Novack et al. Unlike some previous interventions in math classrooms [40], Owlet does not focus on a purely digital space. We aim to combine digital and physical, building off prior research done on each of these aspects.

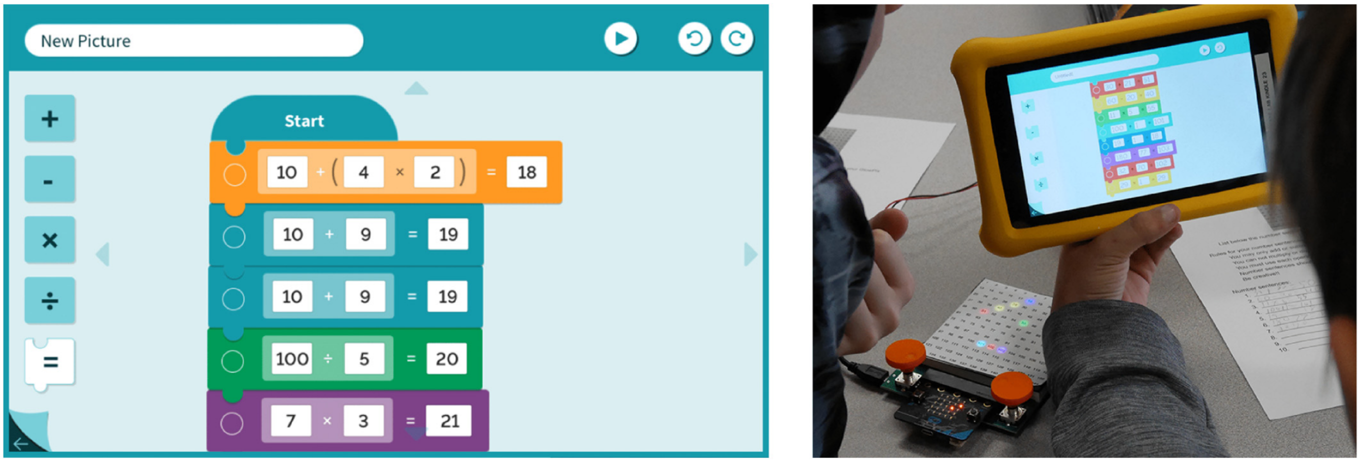
3. Owlet program overview

The design of our tangible manipulatives was informed by: (1) our theoretical framework, (2) reviewing existing US math curricula in grades K-5, and (3) four focus group workshops with teachers across those grade levels. The design process resulted in the Owlet math program, which utilizes two tangible interfaces, GlowBoard and CubeTower; four tablet apps, GlowGrid, GlowPix, Fractions, and CubeApp; and supporting teacher materials. The four apps and their associated manipulatives are shown in Figs. 1, 2, 3, and 4.

Our curricular review consisted of three of the most common curricula used in the United States and the US math common core standards. We found focuses on arithmetic, fractions, and place value as shown in Fig. 5 [41].

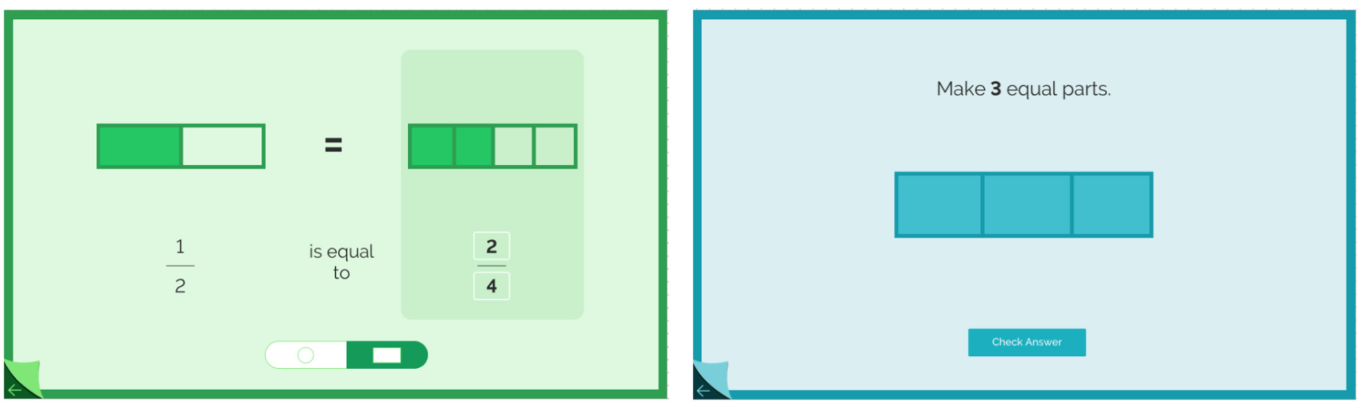
Based on this review and the initial teacher focus group session, we set six Design Goals for the program. In particular, the program should:

DG1. Provide a tangible manipulative interface that students can intuitively use to visualize and explore math concepts.



(a) (b)

Fig. 2. The GlowPix app (2a) supports arithmetic fluency through block based programming. The student creates a block, changes the color, fills out the equation, solves the equation and the corresponding number lights up on the GlowBoard (bottom 2b).



(a) (b)

Fig. 3. The Fractions app supports fraction math concepts using the dials of the GlowBoard as an input.

(a) (b)

Fig. 4. CubeApp supports place value understanding through the utilization of CubeTower, a manipulative with sensors detecting stacks of plastic cubes representing each of three place values. As students place cubes in the CubeTower (4b), the CubeApp (4a) displays the number created. In this example, the Explore activity shows a visual of the number using sets of one, ten, or a hundred squares.

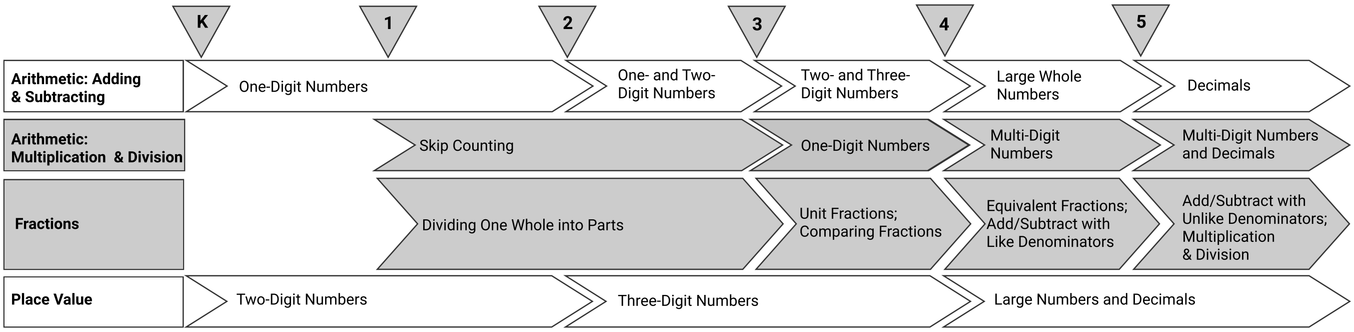


Fig. 5. This figure shows an abridged summary of the topics covered by three curricula in the US as informed by a curricular review [41].

DG2. Support problem-solving at a range of levels of complexity across grades K-5.
DG3. Encourage attention to math concepts while eliminating distracting details.[1]
DG4. Support best practices for teaching mathematics, such as supporting multiple problem-solving strategies.
DG5. Be compatible with any primary math curriculum and integrate into math instruction in a variety of ways.
DG6. Provide a solution at an affordable price point.

We used these design goals to develop the two prototype tangible manipulatives with teachers during the remaining three focus group sessions. Both manipulatives communicate with the tablet via Bluetooth Low Energy. To keep the Bluetooth pairing process accessible to a wide age range, we developed a color coding algorithm allowing students to pair their device by matching LED lights on the manipulative with corresponding colors displayed in the app.

3.1. GlowBoard

The GlowBoard is an array of LEDs created from 144 tri-color LEDs arranged in 12 columns and 12 rows with two rotary dials for controlling the illuminated LEDs and providing input to the apps. The LED matrix provides an interactive representation of traditional tools such as a number chart or a multiplication table (up to 12 $\times$ 12). We provided teachers with overlays of different number charts which fit over the LED matrix, for example, an overlay that supports easier base-10 operations by obscuring two columns. We made minor revisions to the GlowBoard hardware between pilot phases switching out single turn dials with continuous rotation dials for easier manipulation (Fig. 1).

The GlowBoard pairs with three (out of four total) unique tablet apps, to address different aspects of the curriculum such as arithmetic and fractions.

3.1.1. GlowGrid

The GlowGrid application focuses on providing instructional support and practice for arithmetic operations. Students choose one of five sections: *Addition, Subtraction, Multiplication, Division, or Division with Remainders*. For each operation, the GlowGrid app includes up to five types of activities:

1. *Explore* - Students turn the GlowBoard dials to set the terms of an arithmetic equation. The tablet and GlowBoard show the students the equation that they are modeling.
2. *Make* - The app challenges students to use the operation to make a target number in different ways, e.g., Make 24 in 4 different ways. The students model the equation using the GlowBoard to complete the challenge in the app.
3. *Solve* - The app challenges students to solve arithmetic problems with the GlowBoard.
4. *Find* - The app challenges students to solve missing number problems (e.g., $4 + \_\_ = 10$) with the GlowBoard.
5. *Create* - Students can create their own problems and use the GlowBoard to demonstrate the solution.

In each activity, students use the dials on the GlowBoard as controls for the GlowGrid app. The LEDs on the GlowBoard are used as a visual representation of the numerals and arithmetic shown in the GlowGrid app. These dials also control an accompanying equation and visualization on the tablet screen. The dial controls and LED representations are different for each operation, as shown in Fig. 6. Some of the operations have levels that scale difficulty, for example limiting sums to 20 in *Addition* level 1 for younger students.

3.1.2. GlowPix

GlowPix supports students' arithmetic operations fluency through the use of Scratch-inspired block-based programming of different mathematical equations [43]. The student writes an equation in each block of the program. The block lights up the LED on the GlowBoard corresponding to the equation's solution, creating a pixel-art picture. GlowPix supports a variety of student abilities with different levels, explained below.

1. *Level 1* - Equations in level 1 are limited to addition or subtraction by 1 or 10. Each equation starts with the solution from the previous block in the program. The student completes the equation by choosing to add (or subtract) 1 or 10 to that number (Fig. 7). The GlowBoard is restricted to the numbers 1–120 to align with addition and subtraction charts.
2. *Level 2* - Level 2 is the same as level 1 but, students can add and subtract by any number.
3. *Level 3* - Each equation is independent of the previous equation in the program and can contain up to four numbers. Operations are limited to addition and subtraction, and LEDs are numbered between 1 and 120.
4. *Level 4* - Students write equations for addition, subtraction, multiplication, or division. Each equation is independent of the previous equation in the program. The full GlowBoard (numbers 1–144) is utilized to allow for larger pictures.
5. *Level 5* - Students can create complex equations with parentheses and as many operators as they wish (Fig. 8). Otherwise, Level 5 is the same as Level 4.

[1] According to Kaminsky et al. since extraneous information can compete for a user's attention, educational material should maximize the likelihood of attending to relational structure and minimize the likelihood of diverting attention primarily to the superficial [42].

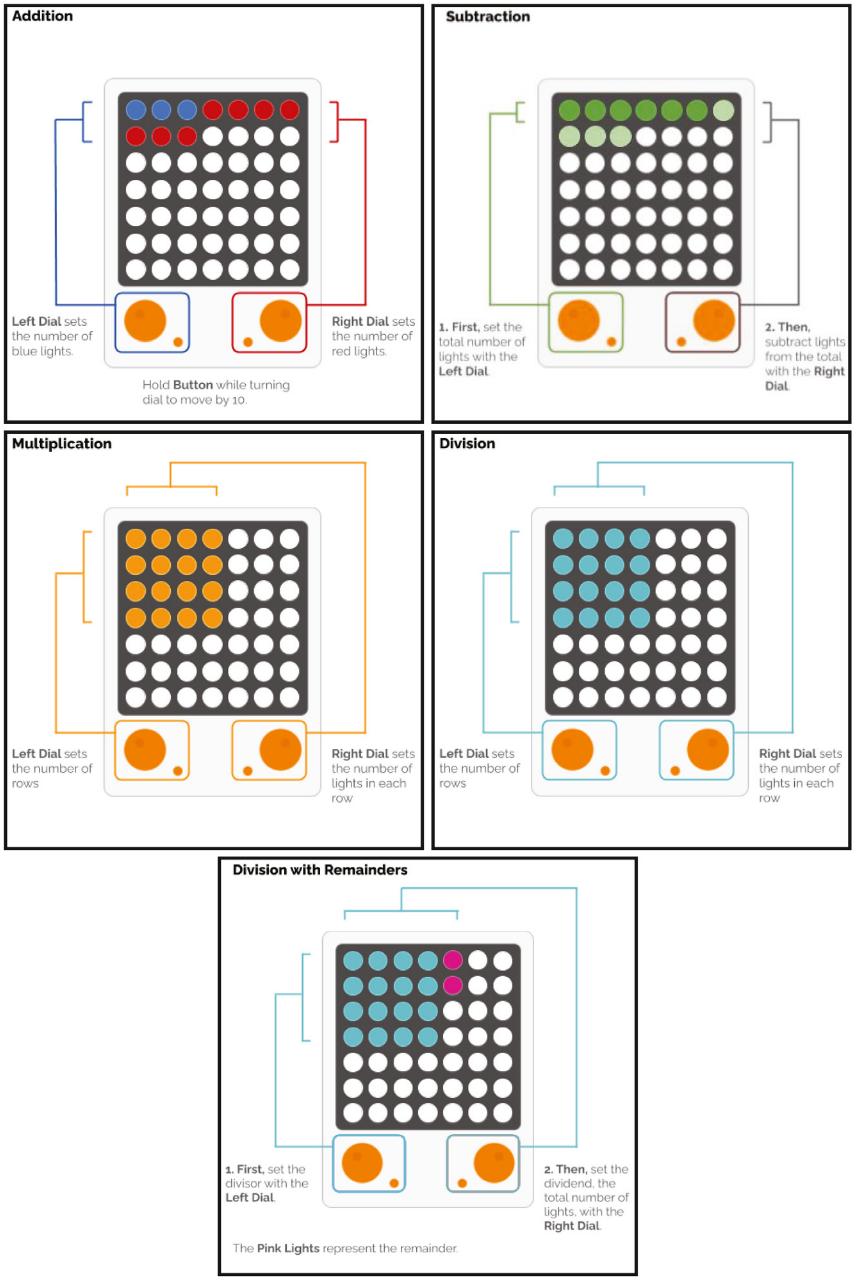


Fig. 6. Five mathematical operations are supported in GlowGrid. The GlowBoard dials control the lighting of the LEDs differently based on the current operation.

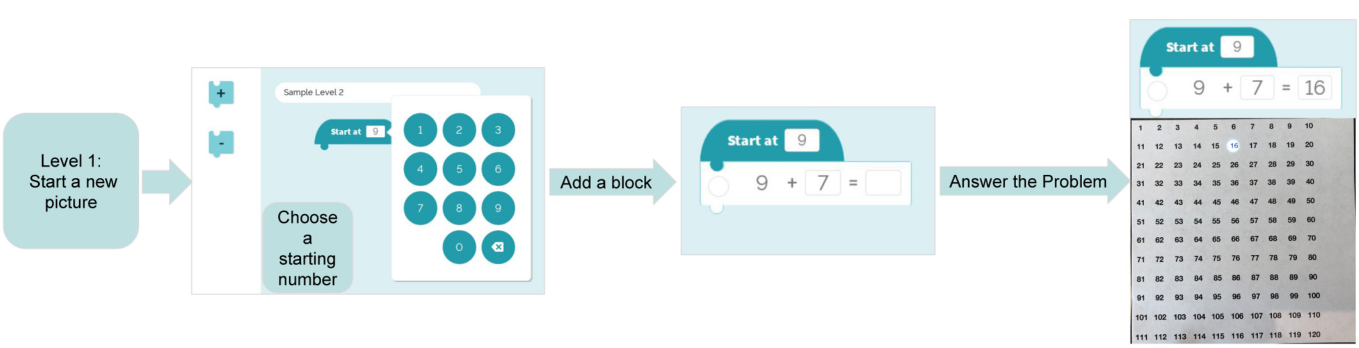


Fig. 7. The workflow of GlowPix in level 1.

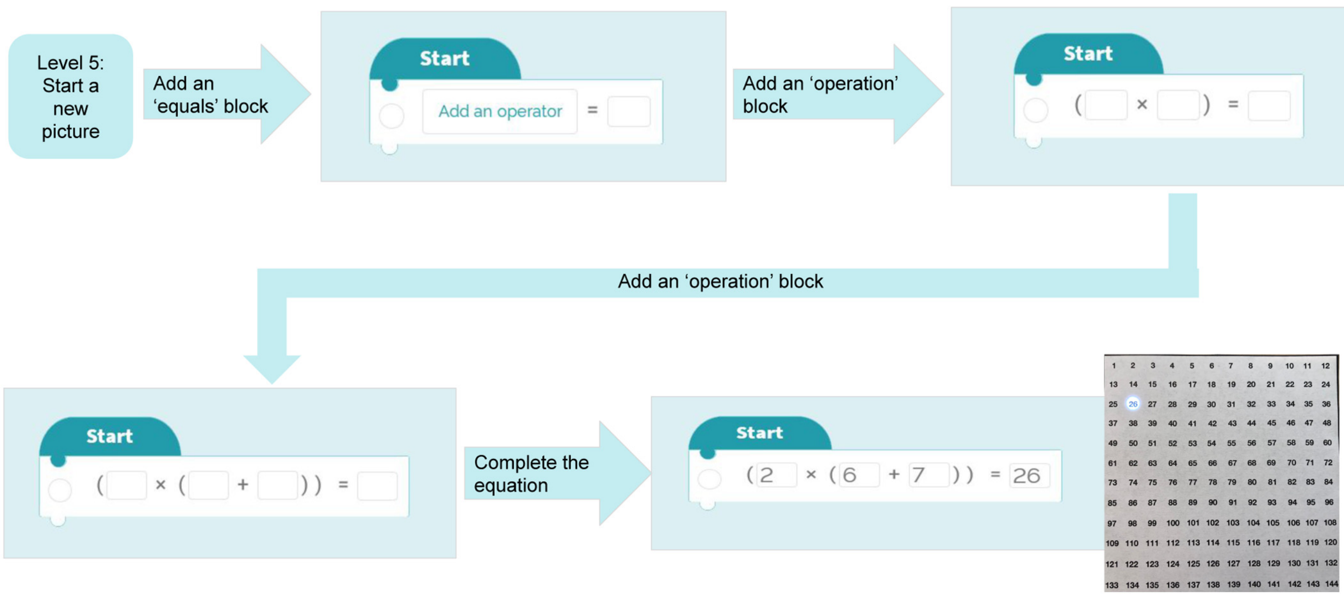


Fig. 8. The workflow of GlowPix in level 5.

3.1.3. Fractions

Fractions has three main sections: *Foundations, Intro, and Equivalence*. Throughout the Fractions app, fraction denominators are limited to 24ths to ensure the on-screen representations remain visible and countable by students.

Foundations is designed to support the foundational skill of recognizing the parts of a whole. In *Foundations*, the left dial on the GlowBoard controls the number of sections each shape is broken into (i.e., the denominator), while the right dial has no control over the app display (as the concept of the numerator is not yet introduced to students).

Intro supports students as they are being introduced to fractions, and *Equivalence* supports students learning equivalent fractions. The *Intro* and *Equivalence* parts of the app use the right dial of the GlowBoard to control the numerator of the fraction on the screen and the left dial to control the denominator. Each of these sections has up to four types of activities, described below.

1. *Explore* - For *Foundations*, students change the number of parts a selected shape is broken into. For *Intro*, students select a circle or bar visualization and control both the numerator and denominator. For *Equivalence*, students create two fractions and the app automatically displays the appropriate comparison.
2. *Make* - Students are prompted to make a number of parts (*Foundations*), a fraction (*Intro*), or an equivalent fraction (*Equivalence*).
3. *Build* - (Available in *Intro* only.) Students are prompted to build a variable number of fractions that are equivalent to a provided fraction.
4. *Compare* - (Available in *Intro* only.) Students are prompted to first make two fractions, then compare them.

3.2. CubeTower

CubeTower has three columns of 9 IR emitter/receiver pairs spaced out on a PCB board to align with plastic cubes as they are stacked on top of each other. These cubes are unlabeled and are concretely countable. A frame around the board aids users in aligning the stacks of cubes with the board's sensors, enabling the system to accurately determine the number of cubes within each column. A single CubeTower can model numbers up to three digits, and two CubeTowers used together can model numbers up to six digits.

3.2.1. CubeApp

We piloted the CubeTower with one tablet app, called CubeApp. We designed CubeApp to support student understanding of the concept of place value. The app supports six place value modes: *Tens, Hundreds, Thousands, Money, Decimals Less Than One, and Decimals Greater Than One*.

For each mode, there are five activities:

1. *Explore* - Students create numbers by stacking blocks in the CubeTower and observing how the number on the tablet screen changes. For example, in the Hundreds section, users can construct numbers from 0 to 999 using three place value columns.
2. *Make* - Students are prompted to make a specific number by placing cubes in the CubeTower. The prompt may be written in a variety of forms (e.g. 47, forty-seven, 40 + 7).
3. *Build* - Students solve number puzzles. For example, students might be asked to build a number between 200 and 300 using 5 cubes.
4. *Compare* - Students create a number in CubeTower and then compare it to another number generated by the app (greater than, less than, or equal to).
5. *Round* - Students use the CubeTower to round a number to a specified place value.

We designed this interface to assist students building connections between countable objects and numeric representations.

3.3. Teacher supports

In addition to the interface tools for students, the Owlet Program includes materials to support teachers in their use and integration of Owlet into their classroom. Before each pilot test, we provided teachers four to eight hours of professional development on the usage of the Owlet tangible manipulatives and the integration of the Owlet program into their curriculum and classrooms. This professional development also included activities where teachers detailed goals they had for their lessons.

We also provided teachers with written documentation. Phase one teachers received a manual on Owlet use and example

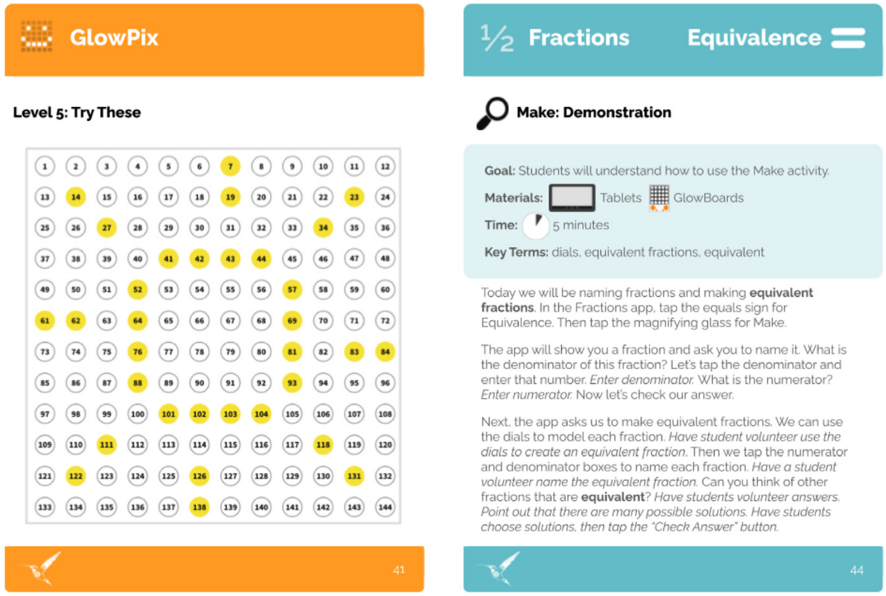


Fig. 9. The card on the left is an example of a GlowPix 'try these' card geared toward level 5 of GlowPix intended as practice for students. The card on the right details a sample demonstration of the make activity in the Equivalence section of the Fractions app.

lessons. Phase two teachers received a set of informational cards covering a variety of topics, with cards for each tangible manipulative and app. The tangible manipulative cards walk teachers through topics such as setting up, using, connecting, and charging the devices; and app alerts given by the devices. The app cards give information on working in pairs, activity demonstrations and discussions, teaching tips, teacher settings, saving work, teacher reports, available levels, and specific app usage information (Fig. 9).

Additionally there are cards for the students including 'how to play' and task cards for each activity, cards on saving their work, connecting to and using the GlowBoard, instructions for dial function within GlowGrid, and 'try these' sample picture cards for GlowPix (Fig. 9).

4. Study design

Testing took place in two phases. Testing phase 1 in Spring 2019 tested the GlowBoard with the GlowGrid app and the CubeTower with the CubeApp. Testing phase 2 in Fall 2019 tested an updated prototype of the GlowBoard with the GlowPix and Fractions apps (see Section 3). For all testing, participants gave their consent by signing a consent form reviewed by an institutional review board.

During phase 1, we used design-based research methods to conduct six in-school tests, one per grade K through 5th (students ages 5 to 11, N = 106). The test schools included two urban schools, one suburban school, and one rural school. Before the test, teacher participants (N = 6) underwent professional development and were provided with the previously discussed supporting materials. The classroom tests were each between two and five days in length (Table 1).

During phase 2, we conducted four in-school tests, one in second grade, one in third grade, and two in fourth grade (students ages 7 to 11, N = 79). These grades were selected to test the target grade levels of the Fractions and GlowPix apps. The test schools included one urban school and two rural schools; one rural school hosted two test classes. Before the tests, all teachers (N = 4) participated in professional development and were provided with updated supporting materials, described in (Section 3.3). Each teacher was assigned one of the two apps and was given the option of testing the additional app. A summary of testing time is provided in Table 1.

Table 1
A summary of the total time spent on each app across all classrooms by grade.

Grade level	Phase 1: Spring 2019		Phase 2: Fall 2019	
	GlowGrid	CubeApp	GlowPix	Fractions
K	135 min	90 min	–	–
1	135 min	90 min	–	–
2	120 min	165 min	120 min	60 min
3	180 min	60 min	40 min	90 min
4	10–55 min[a]	60 min	120 min	180 min
5	90 min	90 min	–	–

[a]Due to an oversight this pilot was conducted with a single GlowBoard and tablet for the first 45 min, then conducted as normal for 10 min.

As part of our investigation of RQ2 and RQ3, we tested Owlet in a variety of grade levels to align with DG2. In Phase 1, we elected to pilot with one class at each grade level in order to identify initial design issues and investigate teacher-usage for our target grade range. We limited our phase 2 testing to grades 2 through 4 to match the curricular content of the Fractions and GlowPix apps. The study sample size and number of classrooms were limited to match the maturity of the program at this stage. We focused on detailed qualitative observations instead of pursuing a six-group comparison study with a larger sample size. Given that some grade levels were represented by a single classroom sample, our analysis does not include age group comparisons as such observed differences could be attributed to individual educator or school differences.

For all tests, at least two researchers observed each class session. Researchers asked participating students what they thought of the tools, observed how the teacher was integrating the tool into their curriculum, and documented what troubleshooting was necessary for the tool to be used successfully. Researchers recorded their observations as semi-structured field notes focused on the documentation of teacher instruction, student interactions with the tools and with teachers and peers, student understanding, tool function, and quotes from students and teachers. We elected to not record video of the classroom test at this stage of pilot testing because our focus was on general classroom

Table 2
The Cohen's Kappa for observation subthemes contained in each theme.

Topic	Cohen's Kappa
General instruction	0.80
Math instruction	0.80
Math concepts	0.76
UX/UI	0.74
Software limitations	0.84
Hardware	1.00
Reactions	0.74

integration and strict restrictions surrounding video recording would have led to a smaller, less diverse sample of participants.

At the conclusion of the tests, teachers were asked to complete a roughly 30-min interview. During the interview, teachers spoke about student engagement (e.g., How was student engagement during the Fractions unit as compared to engagement during your traditional curriculum?), project successes (e.g., How well did you feel the project helped you meet relevant learning standards? Compared to the curriculum you used previously, would you say student learning was less than before, about the same as before, or greater than before?), and areas of opportunity and improvement (e.g., What did students struggle with the most? What can we do to make the LED Array manipulative better?).

4.1. Thematic analysis

Following each set of observations, the research team held debriefing meetings to review field notes, discuss themes in observations, and compile collective notes by grade. If two researchers observed and noted the same incident, these were summarized into a single note. After all tests were complete, the team analyzed the collective notes which were classified by seven broad themes: (1) General Instruction, (2) Math Instruction, (3) Math Concepts, (4) UX/UI, (5) Software Limitations, (6) Hardware, and (7) Reactions. After coding, the team reviewed observations within each theme to extract subthemes (described in more detail in Findings). Pairs of researchers then coded by these subthemes. The Cohen's Kappa for coding subthemes contained in each theme can be found in Table 2.

5. Findings

For each of the themes (listed in Table 2), the emergent subthemes (listed in Table 3) represent the most commonly observed classroom interactions across multiple grade levels. Our observations and findings corresponding to each subtheme are discussed in detail below. Each prominent subtheme is listed with its connection to the research questions in Table 3.

5.1. General Instruction

The General Instruction theme encompassed teacher pedagogical moves and student–teacher interactions that were not directly related to math concepts. The most common subthemes were *Classroom Activities, Use of the Explore Activity, Dial Instruction*, and *Pair Work*.

GI1. Classroom Activities. Teachers used a number of activities in their classrooms during each pilot phase such as supporting the Fractions and GlowPix apps with worksheets, restricting the types of equations students were allowed to use in the GlowPix app, and reviewing the dial operations for the GlowGrid app. Teachers commonly started lessons with an introduction to the tool or a review of how the tool worked.

Table 3
The subthemes emerging through coding highlight trends in the observation data across classes and ages. Each subtheme provides empirical support for one or more of our research questions, indicated by a ✓.

Subtheme	RQ1	RQ2	RQ3
GI1. Classroom Activities	✓		
GI2. Use of the Explore Activity	✓		
GI3. Dial Instruction		✓	✓
GI4. Pair Work	✓		
MI1. Finding Multiple Strategies and Answers		✓	✓
MI2. Using Traditional Manipulatives to Compare to Owlet Manipulatives	✓		
MI3. Teacher Uses the Levels	✓		
MC1. Successful Use		✓	✓
MC2. Teacher Correction	✓		
MC3. Student Self Correction		✓	
MC4. Students Generating Challenges		✓	
MC5. Teacher Using Manipulative as Assessment	✓		
MC6. Differentiated Instruction	✓		
MC7. Students Choosing Not to Use the Manipulative		✓	✓
UX1. Tools Made Sense			✓
UX2. Tools Did Not Make Sense		✓	✓
SL1. Leaving and Rejoining Exercises		✓	
SL2. Perception of Trends in the App by Students		✓	
HL1. Using Battery Packs Over Plugging In	✓		
R1. Grade Level Appropriateness		✓	✓
R2. Student Engagement		✓	
R3. Liking and Enjoyment		✓	

GI2. Use of the Explore Activity. During introduction, students were allowed to freely explore all aspects of Owlet in use that class session. Subsequent lessons using the same tool and app usually began with a review that used the Explore activity which, due to its flexible nature, allowed teachers to include multiple kinds of problems in their lessons that are not provided by the apps, such as word problems. This is supported by subthemes found in the *Math Concepts* theme.

GI3. Dial Instruction. Because the GlowBoard dials functioned differently for different operations in GlowGrid, teachers spent a non-trivial amount of time providing instruction on how to use them. When teachers did not provide this instruction, we observed more students were confused about how to use the GlowBoard. Teachers spent less time providing instruction on the dial use for the Fractions app where the dials function similarly for all sections.

GI4. Pair Work. When used by pairs of students, the GlowGrid and CubeApp activities were more successful when roles were explicitly defined by the teacher. This informed the more defined sample activity cards created for the Fractions and GlowPix pilots. We again observed that student pairs were more successful when following the strictly defined roles as laid out by the activity cards or by the teacher.

5.2. Math Instruction

The Math Instruction theme included teacher pedagogical moves and student–teacher interactions that were directly related to math concepts. When analyzing the Math Instruction theme, we found the subthemes of *Finding Multiple Strategies and Answers, Using Traditional Manipulatives to Compare to Owlet Manipulatives*, and *Teacher Uses the Levels* were most prominent.

MI1. Finding Multiple Strategies and Answers. Frequently in the pilots, as guided by the instructional materials, teachers highlighted the varying mathematical strategies employed by students. Teachers also suggested multiple strategies to students, especially when students were struggling with a concept.

MI2. Using Traditional Manipulatives to Compare to Owlet Manipulatives. Teachers used familiar manipulatives to help students grasp particular mathematics concepts, draw connections between prior lessons and the new interfaces, or to introduce the Owlet tools. This aligns with prior work on transitional representations in concreteness fading [10]. For example, the second and third grade teachers especially noted that the GlowBoard matched the multiplication chart they used previously. The kindergarten and first grade teachers compared the CubeTower cubes to tens sticks.

MI3. Teacher Uses the Levels. We also observed teachers using the level settings to differentiate their instruction for their students across all apps. In GlowPix specifically, teachers added their own rules to further challenge their students. For example, they limited the number of equations that could use addition or subtraction of 1.

5.3. Math Concepts

The Math Concepts theme included student and teacher comments that were focused on math concepts. The most notable subthemes for the Math Concepts theme included *Successful Use, Teacher Correction, Student Self Correction, Students Generating Challenges, Teacher Using Manipulative as Assessment, Differentiated Instruction*, and *Students Choosing Not to Use the Manipulative*.

MC1. Successful Use. When using the tool, students were generally successful; they were able to use the manipulatives as intended by the designers in order to solve math problems and accomplish tasks as presented in the apps. In fact, 34% of the Math Concepts theme observations were coded as the *Successful Use* subtheme.

MC2. Teacher Correction. The teachers were able to guide students through using the tool itself and also through solving math problems using the relevant tool.

MC3. Student Self Correction. In several instances where students struggled to correctly solve a math problem, they were able to correct themselves or their partner through using the relevant manipulative.

MC4. Students Generating Challenges. There were several instances where students would generate their own problems to challenge themselves or their partners. Notably students did this in the explore activities of GlowGrid and with creating their own pictures in GlowPix.

MC5. Teacher Using Manipulative as Assessment. We observed instances where the teachers used Owlet as a way of gauging their students' math abilities. One teacher was surprised at her students' performance when solving problems using the CubeTower. The tool helped her realize that students did not have as strong a grasp of decimals as she had thought; although they had learned decimals previously, students had trouble transferring their knowledge to the new types of problems presented in CubeApp Build challenges.

MC6. Differentiated Instruction. The manipulatives were able to support students with varying levels of math ability. However, most students struggled with higher level fractions problems presented in the app, also discussed in the UX/UI theme. Finding equivalent fractions was a new topic for the students and their struggle with using the GlowBoard and Fractions app showed that the app did not provide enough differentiated instruction to support students at the entry level of these concepts.

MC7. Students Choosing Not to Use the Manipulative. We observed situations where the manipulatives were and were not helpful to students in solving problems. Specifically when using GlowGrid, it seemed that students did not need the manipulative if they felt the problem was too easy or already knew a strategy for solving an arithmetic problem. For example, students chose to use finger counting or mental math to support their problem solving instead of using the manipulative. For the Fractions and GlowPix apps however, students were required by the app to use the GlowBoard as a form of input or output. In the case of GlowPix, students would use worksheets or scratch paper to work out solutions to the more complex equations.

5.4. UX/UI

The UX/UI theme included observations on student and teacher interactions using and making sense of the system. We observed many unique benefits and challenges regarding student use of the interface. Therefore, many of the observations were coded by the level of user comprehension or confusion. As such most observations were coded with two main subthemes: *(UX1) Tools Made Sense* and *(UX2) Tools Did Not Make Sense*, our findings below focus on the latter.

Most notably for GlowGrid, we found that the modeling of subtraction on the GlowBoard did not make sense to the students, substantiating observations in the *General Instruction* theme. Throughout the first phase of pilots, we iteratively made minor changes to the function of dials to make the interaction more intuitive.

When students first received the tablet and manipulatives, they did not expect them to be connected. However, once the students discovered they were linked, they were excited to use the manipulatives. The students did not encounter issues drawing connections between each manipulative and app and could even frequently troubleshoot if their manipulative lost its bluetooth connection to the app.

As noted in the *Math Concepts* theme, students struggled with finding equivalent fractions. We observed students who did not already understand the mathematical principles behind finding equivalent fractions try to solve the problems by using the pictures on the tablet screen. The app, however, was designed as a visual support of the mathematical solution not as a way to solve the problem visually, making it difficult for students to use this approach to solve the problems.

5.5. Software Limitations

The Software Limitation theme included specific issues with the app software. The two main subthemes in the Software Limitations theme were: *(SL1) Leaving and Rejoining Exercises* and *(SL2) Perception of Trends in the App by Students*. Software bugs were also categorized into this theme, and passed onto developers to inform further iterations of the apps.

In GlowGrid and CubeApp each time a student entered an activity, a new, random problem was generated. We observed students utilizing this to generate a different problem when they deemed their current problem too difficult. In reaction to this observation, the Fractions app saved the current problem and displayed it again upon re-entry to the activity. Students then had to solve the more difficult problems, applying math problem solving strategies.

Throughout the first pilots, students perceived trends in the CubeApp Compare activity. Each problem is randomly generated, allowing for the same kind or solution of a problem to appear sequentially (e.g., by entering zero as the comparison number a student can ensure that the answer is always "less than"). In reaction to this observation, the Fractions app contained a more careful, pseudorandom selection of problems, leading to fewer trends in the generated problems and deeper mathematical engagement from students.

Note that these two subthemes only emerged in the observation notes from the first phase of pilots. The more careful problem selection utilized in the second pilot phase required students to engage more deeply with the math content, aligning with prior work by Cuendet, Jermann, and Dillenbourg [16].

5.6. Hardware Limitations

The Hardware theme included observations primarily highlighting problems the research team saw with current prototypes of both the CubeTower and GlowBoard. These subthemes informed future prototypes of the GlowBoard and the CubeTower, including the updated version of the GlowBoard used in the Fractions and GlowPix pilots. The major theme that was not specific to a single tangible manipulative was *(HL1) Using Battery Packs Over Plugging In*. We observed that the teachers took full advantage of the mobility of the devices afforded by the battery packs. Teachers preferred being able to station student pairs anywhere in the classroom without being restricted by extension cords and available outlets.

5.7. Reactions

The Reactions theme included student and teacher emotional responses to and comments in response to the features of Owlet. The most prominent subthemes for the Reactions theme were *Grade Level Appropriateness, Student Engagement*, and *Liking and Enjoyment*. These subthemes from observations were supported by anecdotal teacher interview data as well.

R1. Grade Level Appropriateness. Teachers across grade levels were able to create lessons targeted toward grade-level standards that integrated both the CubeTower and GlowBoard, indicating that the Owlet program was developmentally appropriate overall. The students as well were able to reflect on the appropriateness of the tools for their grade and others. This topic also appeared in the *Math Concepts* theme under *Differentiated Instruction*.

R2. Student Engagement. The research team observed that students were actively engaged with the apps and manipulatives, a sentiment echoed in teacher interviews. The kindergarten teacher noted, "Their engagement was a little bit more noticeable here [...] they really push themselves to stay engaged through that whole time. That is not always what I experience with math, so it was a definite increase in engagement, and they were putting forth more effort than they have with a traditional program".

R3. Liking and Enjoyment. Perhaps most notably, 42% of all reaction observations communicated that students liked the Owlet program and enjoyed using it in class.

6. Discussion

6.1. Features and processes that teachers leveraged in math instruction

In answering our first research question, we considered how features of the Owlet program are used by teachers in their math instruction. During the professional development sessions, teachers become familiar with Owlet and determined how they would integrate it into their classrooms. They noted different features during the professional development that they would specifically leverage for their classrooms. Throughout the pilots we observed teachers putting these plans into effect.

Teachers greatly enjoyed the open-ended aspects of Owlet because they encouraged students to think more intently about the mathematical concepts. Most teachers chose to use the Explore activity, or a more general exploration of the manipulatives and apps, as a way to introduce them to the students. Teachers liked the open-ended discovery that Explore gave the students. Some teachers chose to continue using the Explore activity in GlowGrid and CubeApp to support types of problems not given in the other activities within these apps.

Teachers used the other activities in GlowGrid, Fractions, and CubeApp, as well as GlowPix as a way to allow students to practice math concepts at their own pace. As students worked solving problems, teachers also utilized the levels available in the different apps to ensure that students were working at the most appropriate challenge rating for them. Students worked in pairs during the pilots, also giving teachers the opportunity to use Owlet to focus on pairwork concepts like sharing and taking turns. This aligns with Smith and Walkington's fourth design principle, and the use of pairwork by teachers provides empirical evidence in support of collaboration. The combination of open-ended and leveled activities gave teachers the flexibility to meet the needs of their students and fit Owlet into their math instruction.

On some occasions, the teachers chose to support Owlet with additional materials. For Fractions and GlowPix, teachers often supported the apps with paper and pencil or worksheets to solve and fill out. For CubeApp and GlowGrid, teachers often had students use the apps, then write their solutions on a whiteboard or paper to reinforce math concepts.

6.2. Mediating processes enabling student learning

For our second research question, we examined the processes and activities in the Owlet program which enhance and enable student learning. Tangible user interfaces hold power to enhance and enable student learning through connections and understandings created directly between the physical controls and the digital representation, mediated through the software [28]. In our model, we hypothesize a similar connection occurring within the minds of the students, as they connect physical controls to mathematical concepts (Fig. 10). Our observation of the students supports these connections providing empirical support of Ullmer's theories.

Through our observations and analysis, we reached conclusions on what made the two tangible devices successful. The manipulatives were designed to be used across multiple grades and math abilities, differing from existing manipulatives. The apps did not guide students to use a specific method to solve problems, allowing them to use strategies they found most comfortable, aligning with our fourth design goal. With the GlowGrid app, this flexibility gave students the option to not use the manipulative at all. As discussed in Section 5.3, students would choose not to use the manipulative if they knew the answer already or had a preferred strategy, allowing them to still gain the benefits of practicing the math concepts.

Some aspects of the apps and manipulatives were confusing for the students and distracted them from focusing on the math concepts. For example, the visual representations of equivalent fractions led them to attempt finding equivalent fractions purely visually, instead of numerically and visually. Another instance

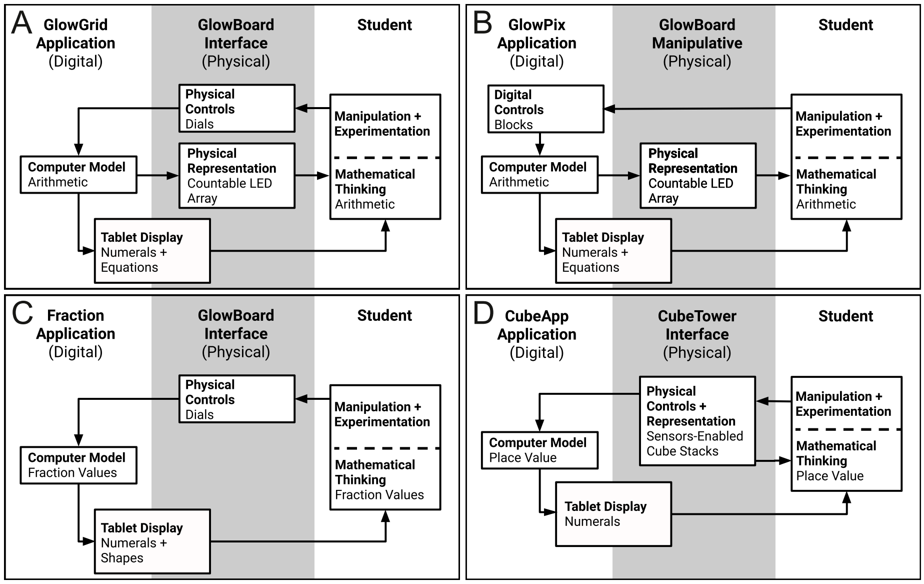


Fig. 10. These system diagrams show the relationships between each of the four apps; GlowGrid (A), GlowPix (B), Fractions (C), and CubeApp (D); and their respective tangible interface, illustrating the differences in physical representation and control of each manipulative.

was the difference in dial use between different operations. This required students to focus first on the use of the tool, but once they became comfortable, allowed them to focus on not only one, but multiple arithmetic operations.

Overall, students were engaged with the Owlet manipulatives and liked using them. This gave them intrinsic motivation to continue using the tool. We observed many instances where students were even motivated to create their own problems and challenge each other. As students worked on these challenges with the manipulatives they were able to correct their own mistakes or their partner's mistakes, learning from each one.

6.3. Interface design trade-off considerations

For our final research question, we considered how interface design decisions and affordances influence students' ability to make connections between physical manipulatives, on-screen representations, and math concepts. The GlowBoard supports Fractions with only a physical control (Fig. 10C), and GlowPix with only a physical representation (Fig. 10B). For GlowGrid, both the physical control (the dials) and the physical representation (LED array) are utilized to support the students; however, they are mediated through the digital computer model (Fig. 10A). The CubeApp, like GlowGrid, is supported by both physical control and representation; however, the physical control (placing blocks) and physical representation (number of blocks) are directly linked by the CubeTower (Fig. 10D). These differences and their influence on the affordances of each interface are described in Table 4.

The Owlet program spans multiple math concepts with apps that require different modes of physical interaction leading us to prioritize flexibility, intuitiveness (needed instruction and cognitive load), and affordability differently for each tangible device. Through user testing, we observed the effect that these differences had on students and teachers in the classroom.

The GlowBoard is designed as a more flexible tool in terms of the types of activities and math concepts that it supports. When paired with different apps it provides opportunities for

Table 4
Characteristics of the GlowBoard and CubeTower and their associated apps.

	GlowBoard			CubeTower
	GlowGrid	GlowPix	Fractions	CubeApp
Mapping between physical controls and mathematical representations	Flexible, multiple mappings			Direct, single mapping
Cost effectiveness (topics covered per cost)	High cost effectiveness; hardware supports multiple domains			Low
Application domain	Arithmetic	Arithmetic	Fractions	Place value
Amount of instruction needed to use the tool	Medium	Medium	Low	Low
Cognitive load of using the system	High/ Medium	Medium	Medium/ Low	Low
Manipulative required to answer question	Not always required	Required	Required	Required

students to solve and create arithmetic equations across multiple operations and create, compare, and solve fraction problems, simultaneously deepening student number sense and mathematical fluency. Therefore, the connections between the controls and representations must be moderated by a mutable computer model to adapt to these different concepts. We hypothesize that the presence of multiple mappings between the GlowBoard and the on-screen representations used in all the GlowBoard related apps causes a decrease in the connection student's make between the physical model and the mathematical concept being modeled, as shown in the first row of Table 4. This is supported through observations regarding dial instruction and students choosing not to use the GlowBoard.

For example, after using *Addition* students in every grade were confused by how to model *Subtraction* with the GlowBoard. The variations in the physical model lead to an increase in cognitive effort needed to use the GlowBoard with GlowGrid. The Fractions app maintains a single input model but contains sections where the right dial is not used. This appeared to require less cognitive effort than GlowGrid (Table 4); teachers spent less time instructing the students on how to use the GlowBoard dials with Fractions than with GlowGrid. The more direct linking reduced the cognitive load, supplying more empirical evidence for reducing cognitive load with multiple representations [44]. In the future, we plan to give complementary operations within GlowGrid more similar dial function in hopes of reducing cognitive load.

The GlowBoard presents an interactive, iconic representation tied to an abstract representation. This connection between representations builds upon the ideas of concreteness fading. In exchange for flexibility the GlowBoard leaves out the physical representations usually seen with concreteness fading, but it still has physical concrete gestures which have been shown to be useful to students when learning math [36].

When looking at design trade-offs between affordability and flexibility vs concrete and constant representation, we chose a design for the GlowBoard that is flexible across multiple operations, concepts, and grade levels, ultimately allowing the design to have more utility in the classroom at a lower cost ratio than CubeTower.

In contrast, we choose to weigh the goals differently in the design of CubeTower and CubeApp, exploring the other direction of the design trade-off. The CubeApp and CubeTower exclusively support understanding of place value. When students first used the CubeTower, most began by immediately putting cubes into the tower, with few needing to be prompted or assisted suggesting a more intuitive interface design. The CubeTower requires less mapping variation between the physical controls and representations because one interface and interaction can represent numbers of different sizes and accommodates the more advanced material covered with older students (e.g. decimals).

The number represented on the CubeTower and reflected in the digital space is a direct result of the number of cubes placed into the tower, as shown in Fig. 10. The connection between physical, digital, and mental is clearer as the CubeTower is manipulated in one consistent way throughout all activities and sections regardless of number size, reflected in Table 4. This work provides empirical support for prior work that found enabling explicit connections between representations more useful for students [10,44]. CubeTower and the CubeApp provide the full range of representations from physical to symbolic as suggested by design principles for embodied math [12]. The physical number of cubes provides a physical representation, the pictures of the place value blocks provide a iconic representation, and the number on screen provides students with a symbolic representation. All three representations are presented and manipulated simultaneously, creating a strong link in the student's mind.

The resulting trade-off is that the CubeTower is a less flexible tool. It enables students to explore only one concept, place value, and, therefore, its utility in the math curriculum is more limited than that of GlowBoard. Covering the range of math topics with single-purpose highly concrete manipulatives would require the costly production of unique manipulatives for each arithmetic operation, fraction magnitudes and more. However, the benefits of CubeTower being specialized for the crucial and foundational concept of place value were visible in our observations of student use. This tension is created by the balance of design goals that eliminate distracting details (DG3) and provide intuitive interfaces (DG1), and goals for flexibility (DG2) and curriculum integration (DG5) all while staying at an affordable price point (DG6). We ultimately addressed this challenge with the creation of two distinct tools with complementary coverage and focus on these objectives.

6.4. Limitations and future work

The findings presented in above sections came from a series of observations in initial pilot studies. These initial observations and interviews with the teachers provided us with qualitative data. Due to the early stages of the Owlet program, quantitative data was not yet collected to confirm the qualitative trends observed in the initial tests. Our observations were made from a single classroom of each grade from K-5, recruited in a specific region and may not reflect the range of outcomes in a larger, more diverse sample. For future work, we hope to test Owlet on larger samples to further enhance the generalizability of our results and to allow analysis by age. Additionally, our pilots focused on analysis of high level classroom integration and teacher uses; in the future we plan to evaluate students experiences and outcomes with video recorded interactions and student outcomes assessments. Moving forward, we plan on enhancing the overall usability of both systems by emphasizing the physical and representational connection as we develop additional apps for the systems to support a varied and holistic approach to mathematics.

7. Conclusion

This paper has two primary contributions. The first contribution is the production of two artifacts, the Glowboard and CubeTower, that support primary math education with tangible interfaces. The second contribution is toward the field's understanding of tangible interfaces, through the evaluation and comparison of two contrasting but similar interfaces. We found that there is an inherent design conflict between flexible utility and intuitiveness. A system where the physical control and the physical representation map to one another through a static, direct, one-to-one mapping is more intuitive for students than systems that can change mappings. A system that changes mappings can increase flexibility and utility across grade levels.

In designing the CubeTower and GlowBoard, we attempted to balance the intuition and flexibility of the tools. Though both systems were engaging and successful, CubeTower was more intuitive, with a stronger relationship between the physical controls, digital and physical representations, and place value concepts. On the other hand, GlowBoard allows for flexibility in its use across grade levels by supporting students while they practice multiple mathematical concepts and arithmetic operations.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: The research and designs presented in this paper are the results of a Small Business Technology Transfer Program grant between researchers and BirdBrain Technologies LLC, an educational technology start-up, with the intention of the commercialization of innovative research. This research may lead to future revenue for the company. Bambi Brewer and Tom Lauwers declare potential conflicts of interest as employees of BirdBrain Technologies LLC. Jennifer Cross, Emily Hamner, and Illah Nourbakhsh declare potential conflicts of interest from the receipt of royalties from BirdBrain Technologies from prior work.

Acknowledgments

The authors would like to acknowledge the contributions of Skye Jensen, Ashley Li, and Parv Shrivastava, who worked to implement the app designs for the pilot studies. We also appreciate the numerous contributions made by all the teachers who participated in the focus groups and pilot studies.

References

[1] National Academy of Sciences and National Academy of Engineering and Institute of Medicine, Rising Above the Gathering Storm, Revisited: Rapidly Approaching Category, Vol. 5, The National Academies Press, Washington, DC, 2010, http://dx.doi.org/10.17226/12999.

[2] G. Kena, W. Hussar, J. McFarland, C. de Brey, L. Musu-Gillette, X. Wang, J. Zhang, A. Rathbun, S. Wilkinson-Flicker, M. Diliberti, A. Barmer, F.B. Mann, E.D. Velez, The condition of education 2016, NCES 2016144, 2016, from http://nces.ed.gov/pubsearch (Retrieved September 17, 2019).

[3] National Research Council, Successful K-12 STEM education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics, National Academies Press, 2011.

[4] A. Bray, B. Tangney, Technology usage in mathematics education research–a systematic review of recent trends, Comput. Educ. 114 (2017) 255–273.

[5] D.S. Niederhauser, T. Stoddart, Teachers' instructional perspectives and use of educational software, Teach. Teach. Educ. 17 (1) (2001) 15–31.

[6] Q. Li, X. Ma, A meta-analysis of the effects of computer technology on school students' mathematics learning, Educ. Psychol. Rev. 22 (3) (2010) 215–243.

[7] P. Moyer, Are we having fun yet? How teachers use manipulatives to teach mathematics, Educ. Stud. Math. 47 (2) (2001) 175–197, http://dx.doi.org/10.1023/A:1014596316942.

[8] P. Moyer, G. Salkind, J. Bolyard, Virtual manipulatives used by K-8 teachers for mathematics instruction: The influence of mathematical, cognitive, and pedagogical fidelity, Contemp. Issues Technol. Teach. Educ. 8 (3) (2008) 202–218.

[9] J.L. Cross, B. Brewer, E. Hamner, L. Zito, S. Speer, M. Tasota, Pilot results of a digital manipulative for elementary mathematics, in: AERA Annual Meeting, San Francisco, CA, 2020, URL http://tinyurl.com/sxgvcuy Conference Cancelled.

[10] S. Suh, M. Lee, E. Law, How do we design for concreteness fading? Survey, general framework, and design dimensions, in: Proceedings of the Interaction Design and Children Conference, in: IDC '20, Association for Computing Machinery, New York, NY, USA, ISBN: 9781450379816, 2020, pp. 581–588, http://dx.doi.org/10.1145/3392063.3394413.

[11] E. Fyfe, M. Nathan, Making "concreteness fading" more concrete as a theory of instruction for promoting transfer, Educational Review (2018) http://dx.doi.org/10.1080/00131911.2018.1424116.

[12] C. Smith, C. Walkington, Four principles for designing embodied mathematics activities, 2020.

[13] A.M. Donovan, E. Fyfe, Connecting concrete objects and abstract symbols promotes children's mathematics learning, 2019, http://dx.doi.org/10.31234/osf.io/ye2j6, PsyArXiv Preprint, unpublished.

[14] E.R. Fyfe, N.M. McNeil, S. Borjas, Benefits of "concreteness fading" for children's mathematics understanding, Learn. Instr. 35 (2015) 104–120, http://dx.doi.org/10.1016/j.learninstruc.2014.10.004.

[15] E.R. Fyfe, N.M. McNeil, J.Y. Son, R.L. Goldstone, Concreteness fading in mathematics and science instruction: a systematic review, Educ. Psychol. Rev. 26 (1) (2014) 9–25, URL http://www.jstor.org/stable/43549781.

[16] S. Cuendet, P. Jermann, P. Dillenbourg, Tangible interfaces: when physical-virtual coupling may be detrimental to learning, in: Proceedings of the 2012 British Computer Society Conference on Human-Computer Interaction, 2012, URL http://infoscience.epfl.ch/record/180395.

[17] S. Ainsworth, The functions of multiple representations, Comput. Educ. 33 (2–3) (1999) 131–152.

[18] M. Suydam, Manipulative materials and achievement, Arith. Teach. 33 (1986) 10–32.

[19] National Council of Teachers of Mathematics, Principles to Action: Ensuring Mathematical Success for All, National Council of Teachers of Mathematics, Reston, VA, 2014.

[20] D. Clements, M. Battisstia, Constructivist learning and teaching, Arith. Teach. 38 (1990) 34–35.

[21] J. Heddens, Bridging the gap between the concrete and the abstract, Arith. Teach. 33 (1986) 14–17.

[22] D.H. Uttal, K. O'Doherty, R. Newland, L.L. Hand, J. DeLoache, Dual representation and the linking of concrete and symbolic representations, Child Dev. Perspect. 3 (3) (2009) 156–159, http://dx.doi.org/10.1111/j.1750-8606.2009.00097.x, URL https://srcd.onlinelibrary.wiley.com/doi/abs/10.1111/j.1750-8606.2009.00097.x.

[23] A. Lillard, Montessori: The Science Behind the Genius, Oxfford University Press, New York, NY, 2005.

[24] E.V. Laski, J.R. Jor'dan, C. Daoust, A.K. Murray, What makes mathematics manipulatives effective? Lessons from cognitive science and Montessori education, SAGE Open 5 (2) (2015) 2158244015589588, http://dx.doi.org/10.1177/2158244015589588.

[25] R. Ambrose, Are we overemphasizing manipulatives in the primary grades to the detriment of girls?, Teach. Child. Math. 9 (1) (2002) 16–21.

[26] D. Ball, Magical hopes: Manipulatives and the reform of math education, Amer. Educ. 16 (2) (1992) 14–18.

[27] D. Willingham, Ask the cognitive scientist: Do manipulatives help students learn?, Amer. Educ. 41 (3) (2017) 25–30.

[28] B. Ullmer, H. Ishii, Emerging frameworks for tangible user interfaces, IBM Syst. J. 39 (3–4) (2000) 915–931, http://dx.doi.org/10.1147/sj.393.0915.

[29] F. Hu, A. Zekelman, M. Horn, F. Judd, Strawbies: Explorations in tangible programming, in: Proceedings of the 14th International Conference on Interaction Design and Children, in: IDC '15, ACM, New York, NY, USA, 2015, pp. 410–413, http://dx.doi.org/10.1145/2771839.2771866.

[30] Tangible Play Inc., Osmo, 2019, URL http://www.playosmo.com/.

[31] S. Schroth, H. Tang, A. Carr-chellman, M. AlQahtani, An exploratory study of osmo tangram and tangram manipulative in an elementary mathematics classroom, J. Educ. Technol. Dev. Exch. (JETDE) 11 (1) (2019).

[32] L.L. Scarlatos, Tangible math, Interact. Technol. Smart Educ. 3 (4) (2006) 293–309.

[33] A.N. Antle, Research opportunities: Embodied child-computer interaction, Int. J. Child-Comput. Interact. 1 (1) (2013) 30–36, http://dx.doi.org/10.1016/j.ijcci.2012.08.001, URL http://www.sciencedirect.com/science/article/pii/S2212868912000025.

[34] M. Giannakos, Z. Papamitsiou, P. Markopoulos, J. Read, J.P. Hourcade, Mapping child-computer interaction research through co-word analysis, International Journal of Child-Computer Interaction 23–24 (2020) 100165, http://dx.doi.org/10.1016/j.ijcci.2020.100165, URL http://www.sciencedirect.com/science/article/pii/S2212868920300015.

[35] G. Baykal, I.V. Alaca, A.E. Yantaç, T. Göksun, A review on complementary natures of tangible user interfaces (TUIs) and early spatial learning, Int. J. Child-Comput. Interact. 16 (2018) 104–113, http://dx.doi.org/10.1016/j.ijcci.2018.01.003, URL http://www.sciencedirect.com/science/article/pii/S2212868916300812.

[36] M.A. Novack, E.L. Congdon, N. Hemani-Lopez, S. Goldin-Meadow, From action to abstraction: Using the hands to learn math, Psychol. Sci. 25 (4) (2014) 903–910.

[37] J. Yu, R. Roque, A review of computational toys and kits for young children, Int. J. Child-Comput. Interact. 21 (2019) 17–36, http://dx.doi.org/10.1016/j.ijcci.2019.04.001, URL http://www.sciencedirect.com/science/article/pii/S2212868918300825.

[38] C. Sylla, C. Coutinho, P. Branco, W. Müller, Investigating the use of digital manipulatives for storytelling in pre-school, Int. J. Child-Comput. Interact. 6 (2015) 39–48, http://dx.doi.org/10.1016/j.ijcci.2015.10.001, URL http://www.sciencedirect.com/science/article/pii/S2212868915000197.

[39] A. Soleimani, D. Herro, K.E. Green, Cyberplayce - a tangible, interactive learning tool fostering children's computational thinking through storytelling, Int. J. Child-Comput. Interact. 20 (2019) 9–23, http://dx.doi.org/10.1016/j.ijcci.2019.01.002, URL http://www.sciencedirect.com/science/article/pii/S2212868917300971.

[40] L. Benton, P. Saunders, I. Kalas, C. Hoyles, R. Noss, Designing for learning mathematics through programming: A case study of pupils engaging with place value, Int. J. Child-Comput. Interact. 16 (2018) 68–76, http://dx.doi.org/10.1016/j.ijcci.2017.12.004, URL http://www.sciencedirect.com/science/article/pii/S2212868917300296.

[41] B. Brewer, E. Hamner, A Review of Primary Math Curricula, White paper, (CMU-RI-TR-20-56) Carnegie Mellon University, Pittsburgh, PA, 2020, URL https://www.ri.cmu.edu/publications/a-review-of-primary-math-curricula.

[42] J.A. Kaminski, V.M. Sloutsky, A. Heckler, Transfer of mathematical knowledge: The portability of generic instantiations, Child Dev. Perspect. 3 (3) (2009) 151–155.

[43] M. Resnick, J. Maloney, A. Monroy-Hernández, N. Rusk, E. Eastmond, K. Brennan, A. Millner, E. Rosenbaum, J. Silver, B. Silverman, et al., Scratch: programming for all, Commun. ACM 52 (11) (2009) 60–67.

[44] S. Ainsworth, Deft: A conceptual framework for considering learning with multiple representations, Learn. Instr. 16 (3) (2006) 183–198.