MAD Learn: An Evidence-based Affordance Framework to Assessing Learning Apps

MAD learn: An evidence-based affordance framework to assessing learning apps

Christothea Herodotou
The Open University
UK
Christothea.herodotou@open.ac.uk

***Abstract*—Existing recommendations about how to select or design mobile applications (apps) for learning have been heavily relied on customer and teacher reviews, designer descriptions, and educational theories. There is a lack of evaluation frameworks that are informed by research evidence of how different children interact and use apps. The first version of an evidence-based framework, coined as *MAD learn*, is presented detailing affordances that hinder or help children's learning, as emerged from relevant studies. To encourage further studies in the field, not only by researchers but also designers and practitioners, a methodological approach to iteratively assess the affordances of mobile apps is also introduced. This is based on (a) visualising the learning design and learning components of a given app and (b) analysing the screen and audio recordings of children's interactions with apps. The proposed approach has been tested with 17 children 5-6 years old who interacted with a maths app. The analysis captured patterns of actual usage, including time spent on different activities, completion rates, communication instances, and number and type of mistakes. Insights revealed that certain design affordances, including instructions, feedback, and help-on-demand, were differently perceived by children, in some cases helping learning while, in others, hindering it.**

***Index terms*—mobile apps, maths learning, children, evidence-based framework, evaluation**

I. Introduction

Mobile applications (apps) are increasingly used by young children in both formal and informal learning settings. In out-of-school contexts in the UK, 49% of young children aged 3-4 are found to use a mobile device to go online and 24% own such a device. By the age of 11, 49% of children own their own device [1]. The significance of mobile technologies for learning in particular, is shown in governmental initiatives aiming to capitalise on the benefits of mobile learning. For example, in the UK, free access to early literacy and language apps were given to disadvantaged families with children aged 2 to 4, as a means to minimise the attainment gap. Young children were shown to fall behind by four months at the age of 5 and good quality mobile apps are expected to minimize that gap [2].

The assessment of the educational quality of mobile apps has extensively relied on reviews by qualified educators, organisations, customers, app developers, and researchers, through the use of specialised rating scales, own experiences of use, and existing learning and Human-Computer Interaction (HCI) theories [3]. In terms of the latter, a top-down approach has been adopted; evaluation frameworks are structured on well-known theories such as cooperative learning, human motivation, and usability design [4] and suggest that good quality apps should support specific learning goals and promote active, engaged, meaningful, and socially interactive learning. Such apps present features such as explicit instruction, repetitive and cumulative training in learning concepts, immediate feedback, challenge and early reward, and individualized, self-paced learning (self-regulation and control) (e.g., [5], [6]).

Although the aforementioned evaluation approaches can inform us about which apps may be good for children, they are based on the expertise and experiences of adults and previous understanding (theories) of children's interactions with information and communication technologies such as computers. In particular, they do not rely on systematic app testing (e.g., observational studies) with end users (in this case children) that would bring to the fore issues such as whether and what children learn from interacting with apps, whether and how educational outcomes align with the designers' intentions, and which design aspects are challenging and potentially hindering the learning process. While existing frameworks are informative in terms of, for example, formulating and testing hypotheses, research with children is needed to capture interactions with mobile apps and inform evaluation approaches. Available studies especially focusing on young children are rather limited, with maths being a domain that has currently attracted less interest compared to other domains such as language literacy [3]. There is still a need for research that isolates and tests specific app characteristics, aligns theory, design and outcome measures, and assessment of varied cognitive and skill-based outcomes [7]. In addition, mobile devices warrant special examination as they present characteristics that distinguish them from other technology-related research (e.g., computers). They have unique features such as a lightweight design, portability, relatively intuitive interface use, and communication features (e.g., [8]). They also support flexible and personalised synchronous and asynchronous learning, anywhere and anytime [9].

In this paper, a technology-affordance approach [10], [11] has been adopted to understanding existing studies and analysing children's interactions with a specific mobile app about maths. The first objective of this paper was to review existing studies about mobile apps and children and produce a first version of an evidence-based evaluation framework (coined as *MAD learn*) detailing affordances that can hinder or help children's learning when interacting with apps. This framework should be seen as a starting point to informing the process of designing or choosing mobile apps for children. The second objective of this paper was to trial an approach to understanding how the design affordances of mobile apps may facilitate or hinder maths learning in young children. Technology affordances are perceived as dynamic, rather than static, socially shaped and co-constructed through interactions between children and mobile apps (See Section II). The approach consisted of two stages: first, visualisations of the app under study were produced in order to identify and record its learning design affordances and how these relate to certain actions by the users. Second, the app was tested with 17 children to identify how design affordances are used or appropriated by end users. This was achieved through screen

This work has been funded by the British Academy Mid-Career fellowship scheme (MD170009).

and audio recordings and analysis of 15-20 minute individual sessions with the app. The mobile app under study was *Moose Math.*

II. Technology Affordance theory

Technology Affordance is a theoretical approach from the field of Human Computer Interaction (HCI) originally proposed by Gibson [10] and elaborated by Norman [12] and Kaptelinin and Nardi [11]. It refers to the actionable possibilities that are enabled or allowed by a tool or technology. Affordances can be 'canonical' [13] that is the conventional and normative actionable possibilities of tools or technologies. For example, the 'canonical affordance' of a cup is to drink liquids, and that of a help button in an app to be pressed and get guidance. Canonical affordances are appropriated by users, for example the use of a cup to hold pencils or throwing items on the help button to produce noises. Tool or technology affordances can be adjusted to situational needs, as defined by users. Affordances either enable or constrain actions, yet user choice or agency is what determines how the technology will be used and how canonical affordances, or affordances defined by technology designers, will be modified and repurposed in practice. Affordances are relative to the user and depend on their ability [14].

In the domain of maths, the theory of affordances has been used to understand children's access and use of Virtual Manipulatives (VM). A meta-analysis of 66 reports [15] identified that five inter-related affordances promote maths learning: (1) focused constraint – focusing student attention on specific aspects of the objects and mathematical processes, thus supporting the process of computation and abstraction, (2) creative variation – allowing students to produce their own representations, thus encouraging experimentation and creativity, (3) simultaneous linking – linking the student with the representation by allowing interaction with graphical, abstract, and dynamic representations, (4) efficient precision – providing students with maths fidelity through precise examples and various copies of dynamic objects, and (5) motivation – students are perceiving the VM as enjoyable, engaging and interesting.

These affordances were replicated in a follow up study [14] focusing explicitly on the affordances-ability relationship. Affordances were used in varied ways (in terms of access and degree of use) depending on children's ability, while the children's ability changed over time influencing their ability to access and the degree to which they used specific affordances. The context, such as characteristics of the task and perceptions of it as being easy or difficult, influenced affordance access. In the same line of work, 95 potentially helping or hindering affordances were identified after 100 children, 3-8 years old, interacted with 18 mobile apps using VM [16]. An example of a hindering affordance is the difficulty to drag a tiny block. An example of a helping affordance is the app affording (providing) the correct answer. Helping affordances were related to positive changes between pre- and post-tests whereas the same affordances were helping for some children and hindering for others.

Similarly, Falloon [17] analysed the screen-recordings of 18 five years old who interacted with 45 mobile apps about literacy, numeracy, thinking and problem solving skills and identified design features that (a) support learning including scaffolds, specific types of feedback and instructions such as apps that resemble the traditional teaching model of a person teaching, the provision of text-to-speech (automatic or as an option) instructions, and corrective feedback (pop-ups or dialogues), (b) inhibit learning including certain types of feedback and instruction, game/learning imbalance, restricted content such as non-corrective feedback, external web-links, and pop-up advertisements, and (c) app-specific features that structure interaction such as access to settings, time limits, levelling up. The absence of any structure led children spend time mostly on the gaming elements of the apps rather than the learning ones.

Overall, there is limited research available showing how children make use of mobile apps and how certain design features facilitate or inhibit learning. In a systematic review of studies with young children, Herodotou [3] summarized the following design features as facilitating learning: (a) the interactivity of mobile apps, as opposed to a video version, was shown to support near, yet not far transfer of learning, (b) narration and highlight functions related to enhanced reading performance, (c) open-ended tasks, variety of representations and varied levels of difficulty supported refinement of maths ideas, (d) extrinsic feedback was not shown to support writing on a mobile device, as opposed to naturally occurring feedback when writing on a paper, and (e) character familiarity was shown to have no effects on learning.

Table 1 (see end of text) introduces a first version of the *Mobile App Design evaluation framework for learning (MAD learn)*. MAD learn summarizes app design features that have shown to relate to learning. These features emerged from relevant research and either facilitate or hinder children's learning with apps. They have been grouped in eight categories as follows: main figure, feedback, instructions, highlighting information, constraints, linking multiple representations, experimentation, and other features such as progression, sounds, and language.

III. Methodology

A. Sample

As part of the project *mEvaluate* funded by the British Academy, five schools were identified through advertisement in teacher association networks and were self-selected to take part in the project. In this paper, data presented are from one of the schools in which 17 children from Year 1 and Reception took part in the study (5-6 years old). The school presented a larger than the national average concentration of disadvantaged students (minority ethnic groups, English as an additional language, free school meal, children in care or adopted, and pupil premium). Mobile devices (tablets and iPads) were provided by the researchers.

B. Process of Data Collection

Participating children had eight sessions with the mobile app of 15-20 minutes each, across four weeks. The purpose of the implementation was to assess the impact of using the app on maths learning through pre- and post-tests. Outcomes from this analysis, reported in a separate publication, shown equally good outcomes between the control group (engaged in maths practice as usual) and the intervention group (using the app). In this paper, the first session of the implementation is analysed. This session was audio- and screen-recorded with support from a researcher, as a means to trial the proposed approach to analysing children's interaction with the app and get an in-depth understanding of how children interacted with the app. Children worked individually (one device per child) with the devices. No guidance or help was provided by the teacher or the researcher, unless technical difficulties inhibited a child's interaction with the app.

C. *Moose Math: The App Under Study*

Moose Math is a drill and practice app mainly targeting remembering/recalling of, for example, simple addition and subtraction tasks and counting. It mostly promotes practicing of already known knowledge rather than development of new maths skills such as unpacking the process of adding quantities or numbers. It is an 'instructive app' [18] with external game-like rewards. At the point of writing, it is rated with 4.5/5 (iOS) and 4.4/5 (100,000+ installs) (Google store). Participating children were asked to interact with specific activities from the app that were deemed suitable to their age, in particular the *Juice Mixer* giving children some numbers and asking them to add the corresponding quantity of fruits in a mixer, *Pets* asking children to match Arabic numbers or number of dots between two pets, and *Pets Bingo* asking children to find the correct answer to a math sentence from a given square of numbers.

The app was selected based on the following: (a) free maths apps available in both the Apple and Google stores, (b) apps not used in previous published work, to expand rather than replicate existing studies, and (c) good rated apps. After reviewing the design of relevant apps, existing apps could be grouped in three main categories: (i) math apps linked to artefacts, (ii) drill and practice apps with external rewards, and (iii) apps that combined gaming and learning elements, for example a racing maths game. In this paper, we present how the design affordances of apps in the second category may relate to learning.

D. *Analysis of Screen Recordings*

The video analysis is based on the screen and audio recordings of the first session during which 17 children interacted with the app. The screen recording captured children's actions on the screen and their decision-making while interacting with the app. The audio recording captured children verbal interactions such as excitement, boredom, communication with peers, the teacher, or the researcher. Video recordings lasted for an average of 20 minutes (M=20, SD= 5.6). To document the video analysis in a systematic manner, the design affordances of the app were visualized following the Activity Theory framework for analyzing serious games (Carvalho et al. 2015). This analysis depicted the way game and learning elements related to each other and achieved specific learning objectives. This framework helped to understand the possibilities and limitations of the app and informed about the affordances of the app, in particular the app activity sequence, instructions, help, and repetition of tasks.

To document the video analysis, two data analyses rubrics were created. A bottom-up approach was followed to creating these rubrics, that is, three videos were analysed in a sequence noting down different types of information captured by the screen and the audio. The aim was to record all observable actions a child undertook while using the app. These categories were then applied to the analysis of the remaining data, and they were as follows: (a) an individual video and audio analysis rubric with the following information: activity name, level of activity, number of wrong attempts before completing a level, whether the level is completed, description of mistakes per level, hindering affordances, helping affordances, affordances not accessed at all, appropriation of affordances (i.e., unintended by designers uses), oral comments by the child, interaction with teacher and peers. (b) a synthesis rubric in which information from the individual video analysis rubric was combined with demographic information. In particular, the following were calculated and reported per child: (i) performance over time as measured by the following variables: percentage of levels completed (divided by sum of levels attempted), sum of wrong attempts before completing a level, type of in-app mistakes.(ii) affordance usage as measured by number of interaction with in-app affordances, and interaction with the teacher and peers, and (iii) demographics: age, gender, mobile device and usage, maths performance.

IV. Results

Fig. 1 depicts a view of the design of the app under study, in particular the sequence of activity (all Figures can be also found in Appendix 2 https://bit.ly/31XbWZV). The app presents three activities ('Houses') to choose from, i.e., *Moose Juice, Pets, Lost and Found*, each one involving one or two different set of math tasks. A reward (i.e., a decoration for a house) is offered prior to entering the activity for the first time. Instructions are given after selecting a task. Help is available in each task and after touching a purple 'Bird'. Each task can be replayed up to three times. If not completed successfully, the app presents a new task. After completing a number of tasks, the user levels up. A reward (a house decoration) is offered upon completing a level. Fig. 2 adopted from Carvalho et al. [19] details the gaming and learning elements of the app under study. As shown on the table, the main learning elements of the app is to solve a task and receive help. The gaming elements relate to building and decorating a house and are not conceptually relevant to any of the learning elements. A follow-up analysis of the help the app provides (See Fig. 3) showed that this can either be automatic or on demand i.e., the child has to tab the purple 'Bird' to get it.

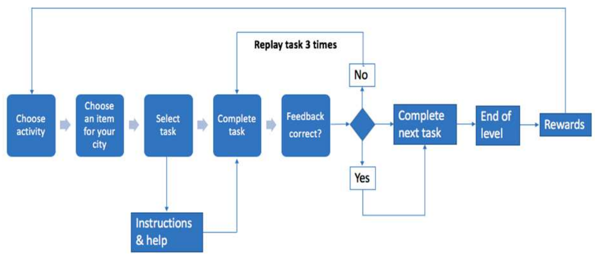

Fig.1. The learning design of the *Moose Math app*

		Choose activity	Choose an item for your city	Instructions/help	Complete task	Rewards
Gaming elements	Actions	Customisation	Choose	Obtain help	Matching	Performance evaluation
	Tools	Activity	Object	Oral or visual instructions	Fruits, bird, drink, oral and visual messages	Rewards
	Goals	Choose activity	Decorate	Learn interface	Solve task	Maximise performance
Learning elements	Actions			Observe	Repetition; recover from errors	
	Tools			Tips/help	Challenge	
	Goals			Provide guidance	Counting or addition or subtraction	

Fig 2. The game and learning components of the *Moose Math* app

Automatic, oral instructions	Optional instructions (help-on-demand)	
	ORAL HELP	VISUAL HELP
Drag ingredients from the fridge to the blender. Put 3 oranges into the blender.		
1st wrong attempt: Let's try again.	Put two more in the blender. I am always here if you get stuck.	
2nd wrong attempt: Let's try again.	Put 2 more Or Remove one	
3rd wrong attempt: The app moves to the next task.		
Correct answer: Looks delicious, choose a cup. Press the cup or tab the recipe book to make another drink.		

Fig 3. Instructions and help provided by the *Moose Math* app

In Appendix 1 (https://bit.ly/39RYYBu), a detailed analysis of the performance of each child, the affordances they did or did not access, communication instances, and demographic information (age, gender, mobile device usage, maths performance) is provided. In this section, an overview of this analysis is presented. In terms of previous usage of a mobile device, nine of the children had no or limited access to a mobile device, while eight of them had frequent or everyday use. The great majority of children (n=14) had an average maths performance, as assessed by the class teacher.

In relation to their interaction with the app, the majority of the children (n=10/17) had less than 100% level completion indicating that they could not complete at least some of the levels they attempted. Seven children (n=7/17) completed all the levels they attempted, yet the majority of them had one or two wrong attempts. Two children only had five wrong attempts before managing to complete all the levels successfully. The overall number of wrong attempts recorded when playing the app ranged between zero and 16. One child had no wrong attempts and 100% level completion – she has exceeding overall maths performance. One child had 16 wrong attempts and a low level of completion (28%). He made systematic use of mobile devices (5-6 days per week) at home and his performance was average (expected). Similar to other children (e.g., Number 7 and Number 11, Appendix 1), repeated mistakes inhibited level completion. An exploratory correlation analysis between demographics data and usage performance confirmed this observation; the only two variables significantly, yet negatively correlated were the number of mistakes made and the percentage of correct answers. The greater the percentage of correct answers, the less the number of mistakes made ($r=-.787$, $p=.01$). Future studies with a larger sample should focus on examining further how certain demographics such as age, gender and maths ability may explain engagement and learning from maths apps.

Eight children asked for help from a peer or the teacher, some of them more than once. These children were found to be uncertain about next steps and confused asking questions such as "What do I do if I make a mistake?". The most commonly observed mistake was found in the *Mouse Juice* activity, when children added more or less fruits than needed in the mixer (n=12). There may be different explanations to this phenomenon; children may have not noticed that numbers indicated the amount of fruits required, have not paid attention to the instructions, or they did not know how to count correctly. Other mistakes observed were: the addition of random fruit quantities in the mixer, filling up all the space in the mixer with fruits without counting what is needed, filling up the app screen with dots, and deleting all given dots and drawing new ones from scratch. These activities could be conceived as forms of appropriation of the app (See Profile 2, below) [17] as children used the app in ways that have rather not been intended by designers and did not align with the purpose of the activity. In terms of the help button, 11 children did not use it at all, while they all had some or many wrong attempts, suggesting that they either did not notice that the purple "bird" when touched was offering help, or they wanted to work the answer out by themselves. In terms of the latter, the audio analysis showed that five of these children asked for help from the teacher or other children multiple times, indicating that they were in need for some help and that they had not realised that help could be found in the app. In terms of the children who made use of the help button, this was not found to be helpful in the majority of the cases; after accessing it multiple times, the children's performance did not change, as they did not manage to find the correct answer or they requested further help from peers, in addition to accessing the help button.

The interaction of children with the app was further analysed in relation to how affordances were accessed or used and whether they helped or hindered learning, as evidenced in children's subsequent actions and their performance in the activities. The analysis of children's interactions with the app, as captured by the rubrics, enabled the clustering of children based on common characteristics, resulting in four main profiles of affordance use. These profiles summarise how children made use of the app affordances and in particular whether affordances helped or hindered learning. Some children could fit into more than one profiles (See Combination of profiles, below). These are depicted below with an example from each category:

Profile 1: Expected affordance usage. One representative example of this profile is a female child in Reception who used the Moose Juice activity. Her overall performance in maths was average or expected (as judged by the class teacher). She completed correctly 20/21 tasks, with the first attempt. She only did one mistake in the first task, in which she did not pay attention or understand the oral instructions and filled up the entire space of the mixer with fruits. This resulted in the wrong answer. She then corrected the mistake and got all subsequent answers correct. She did not make use of the help button at all and she did not seek help from peers or the teacher. It could be argued that the affordances of the app facilitated or did not hinder in any way her performance.

Profile 2: Unexpected affordance usage. This profile refers to instances of app appropriation or use that is unintended by designers [17]. One representative example of this profile is a female child in Year 1, using the Pets activity. She completed correctly 14/14 tasks, and she had two wrong attempts. One of the tasks asked children to match the number of dots across two pets. The 'child' pet showed three dots and the 'parent' pet showed only one dot, requesting the addition of two more dots. This child in her first attempt added three, instead of two dots, and received feedback that this is the wrong answer. In her second attempt, she deleted the given dot on the 'parent' pet and drew from scratch three new dots.

This is the approach she followed in all other tasks where some dots were given and children were requested to add the missing quantity. She did not make use of the help button (‘Bird’). At some point and after listening to the automated oral instructions, she requested help from another child in order to proceed with the task.

Profile 3: Hindering (or non-helping) affordances. One representative example from this profile is a female child from Reception, using the Moose juice activity. She completed 2 out of 8 tasks, and she had wrong attempts in 7 tasks. In the first 7 tasks, she kept adding wrong quantities in the mixer ignoring or not noticing the number indication. It is noted that the number is flashing the first time the task is completed and that oral instructions point to the number. The automated oral instructions did not help the child modify their approach. The child did so only after the teacher pointed to her the number of fruits next to the mixer. She did not access the app help (‘Bird’), yet sought help from the teacher. In this case, app affordances including the oral instructions, the number indication and the help button did not help the child overcome the difficulty she was facing.

Combination of profiles (Profiles 2 and 3). One representative example of a child that fits into two profiles is a male child from Reception, using the Moose Juice activity. He had 6/21 correct tasks, and he had wrong attempts in 16 tasks. The five tasks he completed correctly asked for only one type of fruit to be added to the mixer. In all other tasks requesting to mix two or more types of fruits together, he was unsuccessful. He was repeating the following mistake: he was adding only one type of fruit, despite the automated oral instructions explaining what it should be done. He accessed the help button (‘Bird’) five times, yet this was not effective as he did not manage to find the correct answer. At three different instances, he deviated from the learning objective of the activity, and “shoot” pieces of fruits on the ‘Bird’, often saying out loud: “'Fire!". Throwing pieces of fruits to the bird could be seen as a form of appropriation or non-expected affordance use (See Profile 2).

V. Discussion

In this paper, a first version of an evidence-based affordance app framework, the *Mobile App Design evaluation framework for learning (MAD learn*) has been introduced. MAD learn presents a number of app affordances that were shown to either hinder or help children’s learning when interacting with learning apps. These affordances have been grouped in eight categories: main figure, feedback, instructions, highlighting information, constraints, linking multiple representations, experimentation, and other features such as progression, sounds, and language. MAD learn should be seen as a starting point for the design of educational apps that are grounded on evidence from research and are considering for how certain app features may enable or inhibit learning progress and achievement. The proposed features should be perceived as dynamic, suggesting that they may be appropriated differently depending on the abilities, previous experiences, age etc. of children. For example, a hindering feature may trigger persistency and contribute to the successful completion of a task, depending on the age of a child [20].

To understand how children may actually perceive and appropriate app features, a methodology has been presented and trialled with 17 children. This is comprised of two steps: a) an analysis of the learning design and the learning elements of an app, thus visualising its features and qualities and b) a followed-up empirical examination of the app and its features with children, which reveals how these features are appropriated by certain children. In this paper, the analysis of 17 screen and audio recording videos provided insights that explain how the design affordances of a maths app may facilitate or hinder completion of a set of maths tasks. Overall, it was revealed that affordances including in-app instructions, task feedback, and help-on-demand may not be perceived in a similar manner by children. In terms of oral instructions about how to complete a task and the task feedback (whether the task is correct or wrong), some children were found to ‘adhere’ to them and were successful in correctly completing the task in hand. Yet, other children did not access, did not understand or did not pay attention to them and used the app in unexpected ways. In terms of the help-on-demand, some children did not notice or choose not to access the help button at all. Instead, they requested help from peers and the teacher in the face of difficulties. Other children did access the help button and received additional oral and visual help from the app, yet this feedback was not effective as it did not help them to modify their learning strategy and find the correct answer.

Children’s interactions with the app were clustered into the following profiles: (a) Profile 1: Expected affordance usage. This profile refers to children who used the app as ‘expected’ or as ‘intended’ by designers. These children had few or no mistakes and completed most of the given tasks they attempted with no major difficulty. (b) Profile 2: Unexpected affordance usage. This profile refers to children who used the app in unexpected ways or children who appropriated the app. Examples of unexpected uses are: the addition of random fruit quantities in the mixer, filling up all the space in the mixer with fruits, filling up the app screen with dots, deleting all given dots and drawing dots from scratch, and ‘firing’ fruits on the ‘bird’. These unexpected uses suggest that children may have not paid attention or understood the in-app oral instructions, or if they understood these, they sought to experiment with the app functionality, suggesting that children would endorse a degree of flexibility in the design of apps. (c) Profile 3: Hindering (or non-helping) affordances. This profile refers to children who faced considerable challenges when completing the in-app tasks. The app instructions, task feedback and help-on-demand were less supportive of their learning as they did not scaffold the children in overcoming difficulties and complete the learning tasks.

The app under study could be described as instructive [18]; it is mostly a drill and practice app, with game-like external rewards. Its design reflects the design of the great majority of learning apps available in the market (see also [21]). The strengths of this type of learning design is the potential to provide immediate feedback and automate simple skills such as addition of small sums through repetition (ibid). What was shown in this study is that one of the advantages of this app, that is the provision of immediate feedback, in a number of cases did not scaffold successfully young children, help them to modify their learning strategy and complete the learning tasks. As shown in Fig. 3, the task feedback (received after completing a task) is verbal and emotional in nature (e.g., Let's try again or Looks delicious) or at a 'self-level' referring to personal evaluations and affect in the form of reinforcement [22]. In contrast, help-on-demand provides feedback at a 'task-level', that is instructions about how to proceed (ibid). These instructions are verbal, written, and graphical (images) (see Fig 3). What it is suggested as the most beneficial form of feedback for supporting learning is elaborative feedback, that is explanations as to why an answer is correct or wrong, and/or cues and suggestions as to how to modify a response [23]. In Moose Math, elaborative feedback is provided in the help-on-demand, yet not in the task feedback, suggesting that the latter could be enhanced by explaining to learners why an answer is correct or wrong. Also, the verbal, written and graphical provision of elaborative feedback in the help-on-demand was not shown to be effective in this study, suggesting that designers should seek to test other ways of providing elaborative feedback to young children. A suggestion could be the provision of visual clues that showcase how to move items on the screen and complete a task as well as verbally describing the actions needed for finding the correct answer. These observations suggest that the specific app, when used by certain children, should be supported by more knowledgeable others (parents, siblings, teachers etc) that could provide additional scaffolding when needed rather than used as a stand-alone app.

In addition, the location and size of interactive elements such as the help-on-demand button may explain the poor performance of some children. It may be the case that the size of the help button was too small for children to notice or placed on the screen in a position that was not prominent enough for children to identify and interact with. In other instances and aligning with existing studies [20], children may have found it hard to move with precision items on the screen, resulting in mistakes in the learning task. It is noted that this was the first time children interacted with the app under study and the lack of familiarity or experience in using the app may explain some of the mistakes observed and why some features were found to hinder learning. As a follow-up of this study, the video recordings of the last session of children's interaction with the app will be analysed and compared in order to identify how familiarity with the app may affect affordances and performance.

Learning of basic maths concepts such as addition, subtraction and multiplication requires engagement with multiple representations at the concrete, representational and abstract level [24]. Learning has shown to be effective when it initiates with solving problems with concrete objects, moves to solving problems with representational drawing (dots, lines etc.) and finally solving problems without any support [25]. In the app under study, children progress through the app by completing levels. A closer inspection of the type of activities in these levels reveals that in, for example, the Moose Juice activity, there is an increase in the level of difficulty between Level 1 and next levels. Level 1 requests addition of one type of fruit in the mixer while in subsequent levels is a mixture of Level 1 tasks and tasks requesting addition of more than one type of fruits in the mixer (e.g., 2 bananas + 3 pears). Similarly, in the Pets activity, Level 1 requests to match the number of dots (graphical representation of dots) or the Arabic representation (e.g., 4). In subsequent levels, there is a mixture of Level 1 tasks and tasks that require to match the number of dots and their colour or some dots are given and children are asked to add the missing ones. Yet, the sequence/ordering of tasks does not consistently progresses from concrete to abstract tasks or from easy to difficult levels. Aligning with existing research on how children learn maths (e.g., [26]) lower levels should have been structured around graphical number representations only, subsequent levels should combine graphics with Arabic numbers and the 'difficult' levels should present tasks with numbers only, thus catering for the transition from representational to abstract levels. Concrete representations could be catered if manipulatives (real objects) or in-app virtual manipulatives were used alongside the app.

VI. Conclusions

This paper flags the need for a systematic examination of the design affordances of mobile apps in order to test and isolate the effects of different design affordances on children's learning and understanding, especially when the apps are used under self-regulated learning conditions and with no support from others. In this paper, the MAD learn framework was introduced as a means to provide evidence-based recommendations about how to design apps that facilitate learning. Also, a methodological approach, presented and trialled with 17 children, provided insights as to how to assess the affordances of apps before shared in the market. Both the MAD learn framework and the proposed methodological approach should be treated as an example of how researchers and designers could examine in a systematic manner the design affordances of their app and enable understanding of whether and how the design can support or hinder children's interactions and task completion.

Insights from the observational study are enriching the *MAD learn* framework in three categories: (i) Feedback – the provision of in-app feedback should be heavily considered as a critical component of app design; this study suggests that sufficient and elaborative feedback is needed to promote learning and support error recovery, (ii) Experimentation - a degree of design flexibility could cater for different learning capabilities and meet children's need for digital exploration, and (iii) Other features in relation to how children learn certain maths concepts. A good understanding of existing evidence about how children learn maths concepts (e.g., concrete to abstract representations to learn addition), what misconceptions and errors they often make when completing tasks without the support of technology, and what strategies should be used to overcome these errors should inform the design of relevant math apps. Issues such as the ordering or sequence of activities (and accompanied level of difficulty) should be informed by existing research evidence.

Overall, a more rigorous approach to designing educational apps should be adopted. Any app that is deemed to be educational should be tested with children prior to its market release to identify how app affordances are perceived by different children, given their developmental competencies and needs. Evaluation studies with children should be incorporated into the lifecycle of an app development as a core component, and ideally managed and delivered by learning experts. As shown in this study, a suitable way of capturing and understanding children's interaction with app affordances is the visualisation of its affordances accompanied by screen and audio recording analysis. Ideally, children's interactions should be captured using mobile learning analytics algorithms, that can enable automatic data collection and more fine grained analysis.

TABLE I. THE MAD LEARN FRAMEWORK

Design Features	MAD learn framework		
	Facilitating/ Hindering	***Related study***	***More details***
1. Main Figure			
Figure/hero with pause screen and timed questions	Facilitating	Falloon, 2013 [17]	A figure/hero communicating objectives before introducing teaching concepts, using examples and modelling and also asking prompting questions.
Character familiarity	Hindering	Schroeder & Kirkorian, 2016 [27]	Character familiarity was shown to have no positive effects on learning.
2. Feedback			
Corrective feedback	Facilitating	Falloon 2013 [17]	Pop-ups or dialogues about where changes should be made or where thinking is required.
Multiple representations of feedback	Facilitating	Moyer-Packenham et al., 2016 [16]	Accuracy feedback in one or more of three ways: visual (e.g., pictures), auditory (e.g., sounds, oral praise) and numerical (coins and score).
Extrinsic feedback on writing tasks	Hindering	Patchan & Puranik, 2016[28]	Extrinsic feedback was shown to not support writing on a mobile device over naturally occurring feedback when writing on a paper.
3. Instructions			
Spoken instructions	Facilitating	Falloon, 2013 [17]	Instructions (or content) are spoken, can be replayed or played automatically at intervals or after an incorrect answer.
Tutorials and Hints	Facilitating	Moyer-Packenham et al., 2016 [16]	Information tutorials and hints that explain the activity or the game features or show what to do next.
Narration	Facilitating	Masataka, 2014 [29]	Narration related to enhanced reading performance.
4. Highlighting information			
Highlighting	Facilitating	Masataka, 2014 [29]	Highlight functions related to enhanced reading performance.
Pointing	Facilitating	Falloon 2013 [17]	Word highlighting and pointing e.g., where the app was reading supporting learning.
5. Constraints			
Restricting certain interactive features	Facilitating	Falloon, 2013 [17]	Embedded constraints, such as restricting children from working with certain field e.g., playing with props, allows for spending prolonged time on such activities (this helps self-management), time limits to playing with such activities, randomly guessing answers before time is up.
Limiting options	Facilitating	Moyer-Packenham et al. 2016; Moyer-Packenham & Westenskow, 2013[15], [16]	Focused constraint e.g., constraining the number of decimals needed to solve a problem helping children focus on what the answer may be
Easy of moving objects	Hindering	Bullock et al., 2017 [20]	Difficulty in moving objects around, hindering the process of sense-making
6. Linking multiple representations			
Multiple representations	Facilitating	Moyer-Packenham & Westenskow, 2013[15]	Simultaneous linking e.g., linking the student with the representation by allowing interaction with graphical, abstract, and dynamic representations.
Linked representations	Facilitating	Moyer-Packenham et al. 2016 [16]	e.g., implicit magnitude link between the size of animals and the size of the numbers on the line, linking words with shapes.
Linked physical actions	Facilitating	Moyer-Packenham et al. 2016[16]	Embodied actions as part of the design of the app (e.g. asked children to trace shapes with their fingers) can link concrete/sensory representations with the symbolic or abstract ones.
Multiple, precise examples	Facilitating	Moyer-Packenham & Westenskow, 2013[15]	Efficient precision – providing students with maths fidelity through precise examples and various copies of dynamic objects.
7. Experimentation			
Opportunities for creating artefacts	Facilitating	Watts et al., 2016[30]	Creative variation – allowing students to produce their own representations, thus encouraging experimentation and creativity.
Open-ended tasks	Facilitating	Watts et al., 2015[30]	Open-ended tasks supported refinement of maths ideas.
Check before responding	Facilitating	Falloon, 2013[17]	Ability to check a response before final submission (trial ideas before making a decision).
Multiple attempts	Facilitating	Moyer-Packenham et al. 2016[16]	Unlimited/multiple attempts to reach a goal allows for improvements in performance and students are less likely to quit.
8. Other features			
Levelling up /progression	Facilitating	Falloon 2013; Watts et al., 2016[17], [30]	Progressive levels providing scaffolded support from e.g., easy to complex shapes
Automated features	Facilitating	Moyer-Packenham et	Game efficiency features that make certain actions automated

Design Features	MAD learn framework		
	Facilitating/ Hindering	*Related study*	*More details*
		al. 2016[16]	or easier such as zooming features.
Weblinks	Hindering	Falloon, 2013[17]	Embedded external web links
Advertisements	Hindering	Falloon, 2013[17]	Pop-up or banner advertisements
Language	Hindering	Falloon, 2013[17]	Culture specific terms, nouns or accents.
Sounds	Hindering	Bullock et al., 2017 [20]	Sound distractions shown to hinder learning.

REFERENCES

[1] Office of Communications, “Children and Parents: Media Use and Attitudes,” *Ofcom*, no. February 2020, p. 220, 2020.

[2] D. of Education, “Disadvantaged families to benefit from free early learning apps,” 2019.

[3] C. Herodotou, “Young children and tablets: A systematic review of effects on learning and development,” *J. Comput. Assist. Learn.*, vol. 34, no. 1, 2018.

[4] C.-Y. Lee and T. S. Cherner, “A comprehensive evaluation rubric for assessing instructional apps,” *J. Inf. Technol. Educ. Res.*, vol. 14, no. 1, pp. 21–53, 2015.

[5] K. Hirsh-pasek, J. M. Zosh, R. Michnick, J. H. Gray, and M. B. Robb, *Putting Education in “ Educational ” Apps : Lessons From the Science of Learning*. 2015.

[6] S. Papadakis, M. Kalogiannakis, and N. Zaranis, “Designing and creating an educational app rubric for preschool teachers,” *Educ. Inf. Technol.*, vol. 22, no. 6, pp. 3147–3165, 2017.

[7] J. M. Zydney and Z. Warner, “Mobile apps for science learning: Review of research,” *Comput. Educ.*, vol. 94, pp. 1–17, 2016.

[8] G. Vavoula and C. Karagiannidis, “Designing mobile learning experiences,” in *PCI 2005*, P. Bozanis and E. N. Houstis, Eds. Springer-Verlag, 2005, pp. 534–544.

[9] Y. Mehdipour and H. Zerehkafi, “Mobile learning for education: Benefits and challenges,” *Int. J. Comput. Eng. Res.*, vol. 3, no. 6, pp. 93–101, 2013.

[10] J. . J. Gibson, “The theory of affordances.,” in *Perceiving, Acting and Knowing.*, R. Shaw and J. Bransford, Eds. Erlbaum, Hillsdale, 1977.

[11] V. Kaptelinin and B. Nardi, “Affordances in HCI: toward a mediated action perspective,” *Proc. 2012 ACM Annu. Conf. Hum. Factors Comput. Syst.*, pp. 967–976, 2012.

[12] D. A. Norman, *The psychology of everyday things.* New York: Basic Books., 1988.

[13] A. Costall and A. Richards, “Canonical Affordances,” vol. III, no. 2, pp. 85–93, 2013.

[14] S. I. Tucker, P. S. Moyer-Packenham, A. Westenskow, and K. E. Jordan, “The Complexity of the Affordance–Ability Relationship When Second-Grade Children Interact with Mathematics Virtual Manipulative Apps,” *Technol. Knowl. Learn.*, vol. 21, no. 3, pp. 341–360, 2016.

[15] P. S. Moyer-Packenham and A. Westenskow, “Effects of Virtual Manipulatives on Student Achievement and Mathematics Learning,” *Int. J. Virtual Pers. Learn. Environ.*, vol. 4, no. 3, pp. 35–50, 2013.

[16] P. S. Moyer-Packenham *et al.*, “The role of affordances in children’s learning performance and efficiency when using virtual manipulative mathematics touch-screen apps,” *Math. Educ. Res. J.*, vol. 28, no. 1, pp. 79–105, 2016.

[17] G. Falloon, “Young students using iPads: App design and content influences on their learning pathways,” *Comput. Educ.*, vol. 68, pp. 505–521, 2013.

[18] K. Highfield and K. Goodwin, “Apps for Mathematics learning: A review of ‘educational’ apps from the iTunes app store,” in *36th annual conference of the Mathematics Educationa Research Group of Australia*, 2013, vol. 0, pp. 378–385.

[19] M. B. Carvalho *et al.*, “Computers & Education An activity theory-based model for serious games analysis and conceptual design,” *Comput. Educ.*, vol. 87, pp. 166–181, 2015.

[20] E. P. Bullock, J. F. Shumway, C. M. Watts, and P. S. Moyer-Packenham, “Affordance Access Matters: Preschool Children’s Learning Progressions While Interacting with Touch-Screen Mathematics Apps,” *Technol. Knowl. Learn.*, vol. 22, no. 3, pp. 485–511, 2017.

[21] L. Kolâs, H. Nordseth, and R. Munkvold, “Learning with educational apps: A qualitative study of the most popular free apps in Norway,” in *2016 15th International Conference on Information Technology Based Higher Education and Training (ITHET)*, 2016, pp. 1–8.

[22] J. Hattie and H. Timperley, “The power of feedback,” *Rev. Educ. Res.*, vol. 44, no. 1, pp. 16–17, 2007.

[23] V. J. Shute, “Focus on formative feedback,” *Rev. Educ. Res.*, vol. 78, no. 1, pp. 153–189, 2008.

[24] M. Volk, M. Cotič, M. Zajc, and A. Istenic Starcic, “Tablet-based cross-curricular maths vs. traditional maths classroom practice for higher-order learning outcomes,” *Comput. Educ.*, vol. 114, pp. 1–23, 2017.

[25] E. C. Bouck, R. Satsangi, and J. Park, “The Concrete–Representational–Abstract Approach for Students With Learning Disabilities: An Evidence-Based Practice Synthesis,” *Remedial Spec. Educ.*, vol. 39, no. 4, pp. 211–228, 2018.

[26] E. C. Bouck, J. Shurr, L. Bassette, J. Park, and A. Whorley, “Adding It Up : Comparing Concrete and App-Based Manipulatives to Support Students With Disabilities With Adding Fractions,” vol. 33, no. 3, pp. 194–206, 2018.

[27] E. L. Schroeder and H. L. Kirkorian, “When Seeing Is Better than Doing : Preschoolers ’ Transfer of STEM Skills Using Touchscreen Games,” vol. 7, no. September, pp. 1–12, 2016.

[28] M. M. Patchan and C. Puranik, “Using tablet computers to teach preschool children to write letters : Exploring the impact of extrinsic and intrinsic feedback feedback,” *Comput. Educ.*, vol. 102, no. August, pp. 128–137, 2016.

[29] N. Masataka, “Development of reading ability is facilitated by intensive exposure to a digital children’s picture book,” *Front. Psychol.*, vol. 5, no. MAY, pp. 5–8, 2014.

[30] C. M. Watts *et al.*, “An examination of children ’ s learning progression shifts while using touch screen virtual manipulative mathematics apps,” *Comput. Human Behav.*, vol. 64, pp. 814–828, 2016.