Chapter 12

Selection of Apps for Teaching Difficult Mathematics Topics: An Instrument to Evaluate Touch-Screen Tablet and Smartphone Mathematics Apps

I.K. Namukasa, G. Gadanidis, V. Sarina, S. Scucuglia and K. Aryee

Abstract Manipulatives—including the more recent touch-screen mobile device apps—belong to a broader network of learning tools. As teachers continue to search for learning materials that aid children to think mathematically, they are faced with a challenge of how to select materials that meet the needs of students. The profusion of virtual learning tools available via the Internet magnifies this challenge. What criteria could teachers use when choosing useful manipulatives? In this chapter, we share an evaluation instrument for teachers to use to evaluate apps. The dimensions of the instrument include: (a) the nature of the curriculum addressed in the app—emergent, adaptable or prescriptive, and relevance to current, high quality curricula—high, medium, low; (b) degree of actions and interactions afforded by the app as a learning tool—constructive, manipulable, or instructive interface; (c) the level of interactivity and range of options offered to the user—multiple or mono, or high, moderate or low; and, (d) the quality of the design features and graphics in the app—rich, high quality or impoverished, poor quality. Using these dimensions, researchers rated the apps on a three-level scale: Levels I, II, and III. Few apps were classified as Level III apps on selected dimensions. This evaluation instrument guides teachers when selecting apps. As well, the evaluation instrument guides developers in going beyond apps that are overly prescriptive, that focus on quizzes, that are text based, and include only surface aspects of using multi-modality in learning, to apps that are more aligned with emergent curricula, that focus also on conceptual understanding, and that utilize multiple, interactive representations of mathematics concepts.

I.K. Namukasa (✉) · G. Gadanidis · K. Aryee
Western University, London, Canada
e-mail: inamukas@uwo.ca

V. Sarina
Pleasant View Junior High School, Toronto, Canada

S. Scucuglia
Universidade Estadual Paulista (São Paulo State University), São Paulo, Brazil

P.S. Moyer-Packenham (ed.), *International Perspectives on Teaching and Learning Mathematics with Virtual Manipulatives*, Mathematics Education in the Digital Era 7,
DOI 10.1007/978-3-319-32718-1_12

Keywords Apps · Evaluation criteria · Integers · Learning tools · Mathematics thinking

12.1 Apps for Mathematics

Teachers continuously access learning materials that promise to assist children to think mathematically. On a lesson-to-lesson basis, teachers are faced with the challenge of how to select materials that best meet their teaching goals. The profusion of virtual learning tools available via the Internet magnifies the challenge of searching for materials. Moyer-Packenham et al. (2015) assert, "An important goal for mathematics education is the design and selection of mathematics 'apps'" (p. 42). Few studies provide educationally robust reviews on apps for mathematics (Larkin 2013, 2014, 2015a, b; Moyer-Packenham et al. 2015). Several books (e.g., Dickens and Churches 2012), web-based resources (e.g., common sense media—commonsensemedia.org, Children's technology review—childrenstech.com/), and articles in magazines offer lists of top apps and some reviews on selected apps. Reviews of apps on the app store or those Internet sources are largely based on information that advertises the apps (Larkin 2013). Few reviews are based on evaluation of the apps. For example, Larkin (2015a) shares a list of the top 20 apps (e.g., transformations), Larkin (2013) shares the top 40 Number Sense and Numeration apps (e.g., I see!! Math 1), and Larkin (no date) provides detailed reviews of 142 math apps at https://docs.google.com/file/d/0Bwd_RKnZbGDqSUtkOHZsTHdsWVE/edit. In this chapter, we share an instrument for assessing pedagogically useful apps.

Manipulatives—including the more recent touch-screen tablet/smartphone applications—belong to a broader network of learning tools. In this chapter, we refer to touch-screen tablets and smartphones as touch-screen mobile devices. The work of Namukasa et al. (2009) explore the complementary role of physical and Information Communication Technology (ICT)-based manipulatives, also referred to as virtual manipulatives. Virtual manipulatives are interactive and dynamic objects (Moyer et al. 2002). Virtual manipulatives can appear on computer screens, touch screens, holographic images, and a variety of technological environments. Apps are computer applications in which virtual manipulatives (and various end-user software) are delivered on touch-screen mobile devices. Several apps are touch-screen versions of computer and Internet-based applications. The choice of a manipulative—whether physical or virtual (i.e., a virtual manipulative on a computer, a digital board, or a touch-screen mobile device); historical or modern—is complex. It should depend on what is available, what fits the students' culture and expectations, as well as what fits the teacher's system of beliefs (Bartolini and Martignone 2014). Teachers' choices "to use virtual manipulatives in combination with physical manipulatives were influenced by familiarity with similar physical manipulatives" (Moyer-Packenham et al. 2008, p. 215). In addition, even among the same type of manipulatives, these "can be useful or useless depending on the quality of thinking they stimulate" among learners (Bartolini and Martignone 2014, p. 31). According to Hitt (2002), manipulatives are also classified by the specific meaning of a given concept they address (e.g., discrete, linear, or analogical). Educators and teachers need to pay attention to the specific representation categories (e.g., graphic, analytic, or symbolic mathematics) of a given concept that any manipulative—physical or virtual—addresses (Hitt 2002).

In the mathematics education research community, a thread of research focuses on the influence of virtual manipulatives in learning and teaching, on the design modes, and on the quality of these materials (Pepin and Gueudet 2014; Trouche et al. 2013). For a review of literature on the role of mathematics apps, see Calder (2015), Cayton-Hodges et al. (2015), Larkin (2013, 2015a), Moyer-Packenham et al. (2015), Moyer-Packenham and Westenskow (2013), Pelton and Pelton (2012), and Zhang et al. (2015). Some of this work focuses on specific apps: for example, Larkin (2013) focuses on apps for number sense and numeration, Larkin (2015a) on geometry apps, Moyer-Packenham et al. on apps for young children, Zhang et al. on multiplication and division apps.

Several articles (e.g., Peterson 1972; Skip 1990) and online forums (e.g., "negative × negative = positive" at MathForum.org) explore the use of physical, virtual, and visual strategies, among other strategies, for teaching meanings and operations of negative integers. This work builds on the long history of conversations on teaching more difficult concepts such as subtraction, fractions, and integers (e.g., Kamii et al. 2001). More recent conversations focus on how ICT-based technology (e.g., interactive whiteboard, and computer games) could be used to make difficult topics easier to learn.

12.2 Evaluation of Mathematics Apps

What evaluation criteria could teachers use when choosing the most appropriate teaching materials? The increase in the range of ICT-based materials for teaching, coupled with the emergence of a new culture of learning arising with these resources, is creating a need for quality, design, and diffusion criteria, and policies on these resources. Several studies (Calder 2015; Highfield and Goodwin 2013; Larkin 2015a, b; Pepin and Gueudet 2014; Trouche et al. 2013) voice the need for criteria for evaluating ICT-based resources. Pepin and Gueudet (2014) also maintain that the teacher, even in situations where he or she only selects the resources to use, is "a designer of his/her resources" (p. 133). Trouche et al. (2013) assert that new research and policy questions are arising: "Who designs and what do the design processes look like? How to access quality resources?" (p. 771). For Calder (2015), the question is: "What is the [major] motivation of app designers?" (p. 236). To others, the question is about the alignment between a mathematics app and mathematics curriculum for the target group. For example, Larkin (2014) examines the effectiveness of mathematics apps for the Australian curriculum.

A few studies focus on the evaluation of mathematics apps. Some studies utilize qualitative (e.g., Calder 2015; Larkin 2013, 2014, 2015b), and others quantitative, evaluation measures (Larkin 2014). Larkin (2015b) utilized two qualitative measures based on: whether the apps focused on conceptual (deep understanding related to the meaning of mathematics), procedural (following a set of sequential steps to solve a mathematics problem), or declarative (information retrieved from memory without hesitation) knowledge; and their relevance to the Australian curriculum. Of the 142 he fully reviewed, he observed that many of them "were little more than digital flash cards encouraging rote learning." Of the 40 worthwhile apps he evaluated, only 3 apps (Mathemagica, Areas of Rectangles, Maths Galaxy Fun) were exceptional; a majority of apps emphasized declarative or procedural knowledge; only 40 of the 142 apps were "worthwhile mathematical apps to support mathematics learning in primary classrooms" (p. 30); and only 12 apps involved conceptual knowledge. Several of the apps he reviewed were characterized by mismatches: between the mathematics terms in the app name and the mathematics content explored by the apps, between the description of the nature of knowledge (e.g., conceptual understanding) addressed in the app and the actual knowledge explored in the app, between targeted age levels and age levels at which the content of the app is taught in schools, and between the price of an app and the quality of an app.

Among the apps he reviewed, the Number Sense and Numeration strands were dominant. Goodwin and Highfield (2013) found that apps for toddlers, as well as science and literacy apps, dominated their top 10 apps category. Calder (2015), Larkin (2014), and Moyer-Packenham et al. (2015) noted that a variety of educational apps are available for elementary lessons. A majority of the educational apps available are, nonetheless, standalone apps, focusing on one specific content area, and many are drill and practice, only useful for rote learning of declarative and procedural knowledge (Larkin 2013, 2015b). Moyer-Packenham and Westenskow (2013) note the need for research on manipulatives with students beyond Grade 6.

Larkin (2015b) used three quantitative measures in his app evaluations: The Haugland developmental software scale (Haugland 1999); productive pedagogies (Mills et al. 2009); and Learning principles of good games (Gee 2005). The Haugland developmental software scale is based on criteria for evaluating software for young children. It consists of three dimensions: a dimension on the child (e.g., age appropriate, child control, and non-violence), on design (e.g., clear instructions, and technical features), and on learning (expanding complexity, and transformations). Larkin adopted three of the four dimensions of the productive pedagogies identified by Queensland Education (Mills et al. 2009): intellectual quality (e.g., deep understanding, and substantive conversation), supportive classroom environment (e.g., student direction, and academic engagement), and connectedness (e.g., knowledge integration, and background knowledge). The third scale is based on learning principles (e.g., active, interaction, production, customization, agency, challenge and consolidation, critical learning, probing, multiple routes, and transfer) of good video games developed by Gee (2005). Larkin's evaluation scales range from three to ten. Fullan and Donnelly (2015) offer a scale with four ratings for evaluating digital innovations: good, mixed, problematic, and off track. They identify three dimensions including pedagogy, system change (e.g., implementation support, value for money, and potential to diffuse widely), and technology. These studies show the need for instruments for evaluating apps, especially instruments that emerge from studying apps.

Bos (2009b) offers an instrument for determining the degree of fidelity on a three-point scale—low, medium, and high. Bos (2009a), Larkin (2015a), and Moyer-Packenham et al. (2008) study the fidelity—pedagogical, mathematical, and cognitive fidelity—of technology-based learning tools. Bos (2009a, b) builds on the work of Dick (2008) to further elaborate dimensions and degrees of fidelity. To her, mathematical fidelity of a mathematics tool is the tool's degree of conformity to mathematical properties, rules, and conventions of the mathematical content. A tool "should reflect accurately the mathematical characteristics and behavior that the idealized object should have" (Dick, p. 335). Mathematical fidelity is about mathematical accuracy and precision. Cognitive fidelity is about the ability of the tool to lead to learner actions, interactions, and thoughts that embody mathematics concepts or processes, and, potentially, to deeper mathematics actions, interactions, and thoughts. Pedagogical fidelity is about the elements in the tool, such as target-group appropriateness of the content and type of learning activities, that enable students to learn. Pedagogical fidelity is "evidenced… in the organization of the user interface of a technological tool" (p. 334), in features that support valued learning activities and features helpful for learners (Zbiek et al. 2007).

Larkin (2015a) reviewed 53 Geometry apps, evaluating them against the criteria on fidelity, classifying the apps as low-, medium-, or high-fidelity apps in each dimension. He found the apps to score high on pedagogical fidelity and low on cognitive fidelity. Seven (e.g., Coordinate Geometry, Transformations) of the 53 apps scored high on the three fidelities (cognitive, mathematical, pedagogical), and only the top three of these scored consistently high on all three fidelities. Calder (2015) checks to see if a mathematics learning app is appropriate in intended learning and age of users (an aspect of pedagogical fidelity), is applicable to the concepts involved, to enhancing mathematical engagement and thinking (aspects of mathematical fidelity), and whether an app utilizes "visual, sound and movement elements that learners might also find highly engaging" and appealing (an aspect of technical design features) (pp. 243–244).

12.3 Design Features of Mathematics Apps

Major design features identified in the literature on design of learning apps fall under the categories: nature of the app, content, instrumental/interface design, cognitive/intellectual, sociological, and ergonomic aspects (Gadanidis et al. 2004; Sedig et al. 2014). Human computer interactions (HCI) researchers, for instance, argue that well-designed digital tools (also referred to as visualizations or interfaces in HCI literature) are those designed with a deep understanding of cognition. They maintain that the levels of interaction afforded by digital tools vary from those involving minimal cognitive activities to those that involve higher cognitive skills. The levels of interaction afforded also vary from those evoking only physical (touch, feel, see, etc.) actions such as dragging, to interactions such as comparing, to tasks such as identifying and categorizing, and, further, to activities such as problem solving and reasoning. Several key characteristics offered by the digital tools influence higher-order cognitive activities: the range and adjustability of options—the flexibility; number and diversity of interactions; fitness of the interface to the task, to the user, and to the context; and type of transactions ranging from access only, to annotation, modification, construction, and combination of transactions (Sedig et al. 2014).

12.3.1 Digital Learning Objects and Tools

This inquiry on mathematics apps is situated within a larger framework of digital learning tools (LTs) and objects. Gadanidis and Schindler (2006) point out that the term digital learning objects (LOs) involves a variety of designs, from simple digital images or files in pdf format to complex simulations and interactive interfaces. LOs are small interactive programs that are available online and are focused on specific content topics (Gadanidis and Schindler, p. 20). Virtual manipulatives can evolve into mathematical objects (including concepts, procedures, and processes) "when acted upon," patterns perceived, and a new mathematics object emerges to deepen mathematical understanding (Bos 2009b, p. 526). Zbiek et al. (2007) use the term cognitive tools (CTs) to refer to technologies that extend the learning and thinking activities. CTs for mathematics allow the user to act on, compute and externally represent mathematical entities, and involve a variety of designs including simulation, software, micro-world, devices and tool kits. Bos (2011) uses the term *interactive mathematical objects* to refer to the digital learning tools. The tools with a high degree of fidelity enable manipulation in an intuitive way, encourage active participation of the learner, are appropriate for the age level, are mathematically correct, "provide opportunity to construct, test, and revise to understand the patterns and structure the concepts. Manipulating the patterns leads to great depth of understanding" (p. 526).

Maddux et al. (2001) identify two different types of LOs. In *Type I*, the developer determines almost everything that happens on the screen, it affords only "passive user involvement", "a limited repertoire of acceptable responses", "usually aimed at rote memory" and everything that the software is capable of doing can be observed in about 10 min or less (Gadanidis and Schindler 2006, p. 23). In Type II, the user is in charge of what happens on the screen, it affords "active intellectual involvement," the user is in charge of what happens, it is usually aimed at "creative tasks," and many hours are necessary to exhaust what the program is capable of (p. 23). Type II affords a high number of user possible inputs and a high level of interactivity between the user and object. Gadanidis and Schindler recommend LOs involving a hybrid of Type I and Type II. Godwin and Highfield (2013) refer to

Type II as constructive interfaces, with Type I as instructive, and with the manipulable interfaces lying in between. Gadanidis et al. (2004) argue "mathematical investigation, as a pedagogic tool, is not a simple undertaking. Facilitating investigations [by the learners] adds significantly to the complexity of instructional design" (p. 294). According to these researchers:

> Good design becomes possible when mathematics education and human–computer interaction design experts work together, rather than in isolation, taking into account pedagogical goals and interface design principles, and, of course, where there is commitment to test and revise based on feedback from educators in the field. (p. 295)

Bortolossi (2012) observes that factors such as the nature of the mathematical content (mathematical fidelity), pedagogical design (pedagogical fidelity), graphic design, and interface design (technical design features) are fundamental aspects in the production of educational applications. Bortolossi recommends a combination of the best features of several ICT applications to enable, in a rapid-development environment, the creation of low-cost (but richly designed) portable, dynamic, and interactive LTs with a potential for multiple didactic activities. To Fullan and Donnelly (2015), it is important to also evaluate the "underlying digital product model design" (p. 40) along the lines of ease to use, intuitive design, how data are managed, and what experiences it offers the end users.

Commonalities exist among criteria for designing high-quality apps and those for evaluating apps for learning mathematics. We, nonetheless, agree with Larkin (2013) that design criteria for apps may not directly translate to criteria for evaluating high-quality apps for learning mathematics, and with Dick (2008), that design features of learning apps should be selected to serve pedagogical, mathematical, and cognitive principles. Further, Calder (2015) adds that it helps when the motivation of the mathematics app developer is mathematical engagement, rather than profit optimization. On the question raised by Trouche et al. (2013) regarding who designs and what the design processes look like, from our interactions with the app developers on the project, it appears that some app developers are themselves teachers, educators, and educational researchers whose major motivation is pedagogic, or consult, partner with, and seek feedback (or, even, endorsement) on their products from other teachers, educators, and educational researchers. Many of these apps score lowest on cognitive and mathematics fidelity (Larkin 2015a) but higher on pedagogical fidelity. Selected iTunes apps such as Rekenrek by Mathies, Touch Counts by N. Sinclair (an app for Number Sense and Numeration for young children), and MathTappers apps by T. Pelton and Pelton are designed by mathematics educators. Pelton and Pelton (2012) explore the pedagogical practices in the MathTappers apps, some of which support concept development and consolidation of understanding, and others are for fluency building. Larkin (2015a) observes that most educational apps are designed by non-educators and for market reasons. Various publications exist on development and marketing of apps. More work is needed on the design features that influence the usefulness of apps and on how students use the apps.

Trouche et al. (2013) shares a questionnaire with nine different dimensions to measure the usefulness of any Dynamic Geometry Software (DGS), including mathematical content, pedagogical implementation, integration in a curriculum sequence, ergonomic (ease of use) aspects, instrumental content, added value (takes advantage of new possibilities of DGS), potential for use and further modification of the resource. Pepin and Gueudet (2014) illustrate how studies on quality of teaching resources in general have historically focused on mathematical, pedagogical, sociological analyses (such as analysis along the lines of patterns of class of the target audience), or on specific mathematical knowledge, skills and practices. Studies on ICT resources contribute to the dimension on technical, design features including ease of use, quality and uncluttered graphics, and interactivity of the interface (Haugland 1999; Kay and Knaack 2009).

In the early 2000s, when most digital LTs were still designed for use on desktop and laptop computers, Yerushalmy and Ben-Zaken (2004) advocated for manipulatives that could be used on cellphones, since these devices were "an easily available tool that is already part of the culture and daily life… and that is likely to become highly useful for both teachers and students" (p. 3). Mathematics apps for touch-screen mobile devices are now increasingly part of many mathematics classes. Calder asserts:

> The use of mathematics apps, across a range of contexts and age levels, enhanced learning generally, but this was determined to some extent by the appropriateness and applicability of the apps to the particular student, their learning trajectory and the suitability of the app to the particular learning situation. (p. 246)

Basham et al. (2010) voice that "to provide a highly mobile, flexible, efficient, and scalable technology experience for students that could be taken outside of a school's walls… needed to provide students with multiple means for *representation, expression, and engagement*" (p. 340).

12.3.2 Constructive, Manipulable, and Instructive Apps

Goodwin and Highfield (2013) classify digital learning tools by their design features and how the learners' interact with these features into *constructive, manipulable*, and *instructive* apps. The authors define constructive tools as LTs in which learners participate in the generation of representations, tools which are used by the learners as an expressive tool, and tools which offer learners room for higher intellectual engagement, such as for reflection and thinking processes. These tools utilize significant cognitive effort on the part of the learners. Bos (2009a), Larkin (2015b), and Moyer-Packenham et al. (2015) would refer to these as apps with both high cognitive and pedagogical fidelity. Goodwin and Highfield maintain that learning objects that are not primarily constructive may still support learning when they are manipulable.

Manipulable apps may give a predetermined context, use mostly symbolic and iconic images, but still may allow some alteration of representations through user input (i.e., they are likely to evoke moderate to high user engagement). Thus, manipulable apps offer room for experimentation and discovery. Manipulative apps use modifiable graphics. Bos (2009a, b), Larkin (2015b), and Moyer-Packenham et al. (2015) would refer to these as apps with medium cognitive and pedagogical fidelity.

On the other extreme of the spectrum of apps are learning objects that focus only on behavioural learning activities, use symbolic presentations, and present learning in a linear fashion, utilizing repetitive procedural tasks and thus involving very low cognitive investment on the part of the learner. These learning objects focus on the "learner's focus of control over the representations presented on screen" (Goodwin and Highfield 2013, p. 213). Bos (2009a, b), Larkin (2015b), and Moyer-Packenham et al. (2015) would refer to apps that only offer drill activities as apps with low cognitive and low pedagogical fidelity. Zbiek et al. (2007) classified ICT resources such as online textbooks and courses, which were cognitive in nature but only presented information and had no capabilities to offer feedback on the actions of the learner as other resources but not tools.

12.3.3 Emergent, Adaptable, and Prescriptive Apps

Heydon and Wang (2006) assert that curricula paradigms configure the teaching and learning environments in ways that can limit or expand possibilities. Heydon and Wang name three paradigms: prescriptive, adaptable, and emergent. Prescriptive curricula are in line with behavioural psychology views of learning of scripted knowledge. Adaptable curricula involve active interactions and varied roles for the learner to include tailoring of learning activities according to the learner's interests. With emergent curricula learning is co-constructed with others, and learners are also inventors. For Heydon and Wang, constructive apps would support emergent curricula. Manipulable apps would support adaptable curricular. Instructive apps would only support prescriptive curricula.

Students in Goodwin and Highfield's (2013) studies substantially benefited from constructive and manipulative multimedia in terms of depicting multiple representations of concepts and forming sophisticated concept images (Pirie and Kieren 1994). Calder (2015) agrees that the multi-modal representations provide stimulus and novelty "but it is the subsequent thinking that is key to the learning process" (p. 238). The appealing factor is secondary to appropriateness and applicability, to use Calder's terms. Goodwin and Highfield (2013) maintain that constructive apps should not be mistaken to mean "busy" apps, which include extraneous details such as animations, which place unnecessary demands on low-achieving students, and take away from the understanding of mathematics content. Bos (2011) and Calder (2015) observe that distracting animations and colors minimize mathematical engagement.

12.3.3.1 Levels I, II, and III Apps

The app evaluation criteria in this chapter consists of a three-point scale, Level I, II, and III, with Level III a classification of high-quality apps, and four dimensions. It is a qualitative instrument. Each dimension consists of degrees or categories, which lie on a continuum of increasing complexity. That is to say, apps classified as Level III, show the highest degree on a dimension and go beyond the complexity of Level II, and Level II apps go beyond Level I apps. On a given dimension, say the curricula dimension (emergent, adaptable, and prescriptive), it is possible for an app to combine some elements of the adaptable category and a few of the prescriptive category, for example. Gadanidis and Schindler (2006) refer to apps that combine elements from different categories on a dimension as hybrid LOs. Goodwin and Highfield (2013) visualize apps that combine the middle category, manipulable elements, and the top category, constructive elements, as manipulable apps approaching the constructive category. Larkin (2013) found that, whereas some apps fit only in one category on a dimension of forms of mathematical knowledge (conceptual, procedural and declarative), some apps fit in two categories (i.e., they explored both conceptual and procedural knowledge). Classifying apps by levels is in line with reviews aimed at sharing lists of top apps (e.g., Larkin 2014). After Bos (2009b) and Larkin (2015a), we present our evaluation instrument in a chart (see Table 12.1) form to show the varied degrees (or, categories) on each dimension. Level III is the highest score, Level II is the medium score and Level I is the lowest score or most impoverished category, on a dimension. Level I apps are not necessarily off track but apps with characteristics from only the lowest category.

The dimensions of the classification are: (a) the nature of the curriculum addressed in the app—emergent, adaptable or prescriptive, and relevance to current, high quality curricula—high, medium, low; (b) degree of actions and interactions afforded by the app as a learning tool—constructive, manipulable, or instructive interface; (c) the level of interactivity and range of options offered to the user—multiple or mono, or high, moderate or low; and, (d) the quality of the design features and graphics in the app—rich, high quality or impoverished, poor quality. Several of the dimensions and their categories, such as in (a) and (b), emerged from the literature we reviewed, and some, such as in (c) and (d), emerged largely from the process of analyzing the apps. The fifth row is an overall dimension speculating that apps that score high on several dimensions have the potential for intense levels of intellectual/cognitive involvement, those that score high or medium on some dimension would have a limited potential, and those that score consistently low would have the potential for only low intellectual/cognitive involvement. We present details on the dimensions with the evaluation of a selection from the 80 apps we reviewed.

12.4 The Inquiry

The evaluation instrument emerged from a broader inquiry that involved teachers, researchers, and a developer of iOS apps in three contexts. The first was a school context, in which a teacher (who team-taught a unit on integers), in collaboration with the researchers, planned, implemented, and offered feedback during a Grade 7 and 8 integer unit centred on using CTs that enhance pedagogical goals of using

Table 12.1 Classification of middle school apps

Dimension	Level III apps	Level II	Level I
Curriculum dimension			
Address	The *emergent dimension* of curriculum (e.g., building understanding, explaining why, and reflection; this on top of the adaptive dimension)	The *adaptive dimension* of curriculum (e.g., meaning making, on top of the prescriptive dimension)	Only the *prescriptive* dimension of curriculum (e.g., fact masterly)
	Current and high quality curriculum		Dated or no curriculum
Degree of interaction afforded by the App's interface			
Offer	*Modifiable, constructive* interfaces	*Manipulable* interfaces	*Non-interactive, instructive*, access only interfaces
Interactivity and range of options			
Involve	A *high number and diversity of possible user inputs* or selections	A *moderate number and diversity of possible user inputs* or selections	A *very low number and diversity of possible user inputs* or selections
	A *high level of interactivity* between the user and object and with other users *Multiple* interactions	A *moderate level of interactivity* between the user and object *Mono* interactions	The *lowest level of interactivity* between the user and object *Mono* interactions
Technical design aspects			
Utilize	*Multiple* media and *alternative* representations	*Two or three* media and *alternative* representations	*Overly symbolic, linear* interfaces
	Colour, sound, animations, or 3D effects, graphics to focus learning, eliminating those that are superfluous	Colour, sound, animations, and 3D effects graphics to focus learning	*Superfluous and extraneous* details, such as animations, which instead of focusing learning, distract students
Overall, intellectual/cognitive involvement			
Have the potential for	*Intense* (with several opportunities for) *intellectual/cognitive involvement*—also a focus on math connections, understanding, and math extensions	*Limited* (two or three opportunities for) *intellectual involvement*—also a focus on simple application of skills	*Very low* (none or one opportunity) *intellectual/cognitive involvement*—a focus on individual skills and rote learning

manipulatives in teaching. Finding that the materials she had available did not work well for her students, the teacher created the physical version and a virtual version of a manipulative that circumnavigated the errors created by some existing tools. The second context involved work with an industry partner, who provided the researchers with access to the apps his company had developed. The app developer also offered to train team members to design iOS apps for teaching integers. In the third context, the researchers developed an instrument to evaluate randomly selected apps for teaching integers. The results we share in this chapter are from this third context of studying the apps. The initial coding of the apps was based on content, nature of representations used, interactivity level in the apps, nature of the design of the task posed by the app, and relation of the app to other mathematics learning materials. The process was further informed by research literature on the evaluation of apps, resulting in refined categories and other dimensions.

Larkin (2015a) observes that qualitative evaluation instruments are important "for teachers in making decisions about whether or not to use an app" (p. 344). The instrument shared in this chapter could guide teachers when selecting apps that meet the learning needs of their students. As well, it would guide app developers in going beyond apps that are overly prescriptive, focus only on quizzing students, are based on print design, and include only surface aspects of using multi-modality and play in learning, to apps that are more aligned with emergent, high-quality mathematics curricula, apps that focus also on conceptual understanding, and that utilize multiple modes and interactive representations in ways that are central to learning.

12.5 The Apps: How to Tell When an App Is a Useful App

We searched for apps on the desktop iTunes store because more information, including categories of apps, is displayed at the iTunes store as compared to the app store on a phone or tablet. We chose iOS apps because the app developer on the team created iOS apps. As noted by Larkin (2015a, b), locating relevant apps at the app store is difficult by the "sheer number of apps" and "the poor structure of the iTunes app store user interface" (p. 7); the way information on an app is largely based on the developers of the apps and is often inaccurate (e.g., app names on the app store-display names—may differ from names of apps when installed on a device); the way the results are organized and are displayed by icon, only giving the first 100 relevant results; plus the results continually change as new apps are added and old ones are removed or renamed.

We searched for both iPhone and iPad apps. We searched by keywords, including "integer" "negative," and "minus," by a combination of these keywords, such as "negative integer" "negative number", and by other relevant combinations of key words, such as "integer multiplication." The results for iPad apps were, at many times, more than for iPhone apps. Figures 12.1, 12.2 and 12.3 show screenshots of sample results. Because we are aware that app developers place apps in categories and select keywords for their apps based on market analyses, rather than on accuracy of the keywords, we also browsed the apps by categories. In the educational collections apps category, we selected the category of apps for elementary school, as well as apps for middle school and then further selected the category Math Apps. In the category Math Apps for Elementary School, we further narrowed our search by selecting the subcategories Number System/Numbers and

Fig. 12.1 iTunes store apps results for the keyword "negative integer"—iPhone apps

Fig. 12.2 iTunes store apps results for the keywords "negative integer"—iPad apps

Quantity, Early Operations, and Patterns. We also browsed apps under the categories Drill & Practice, Beyond Drill—Strategy, and Beyond Drill—Brain Busters. For middle school apps, we selected the subcategories Pre-algebra & Algebra, and Drill and Practice. Twenty or fewer apps were returned for each of these categories. We did not browse apps for subcategories such as High School Apps, nor the categories such as Geometry and Data, where we did not expect the content of negative integers to be a primary focus. Goodwin and Highfield (2012) found

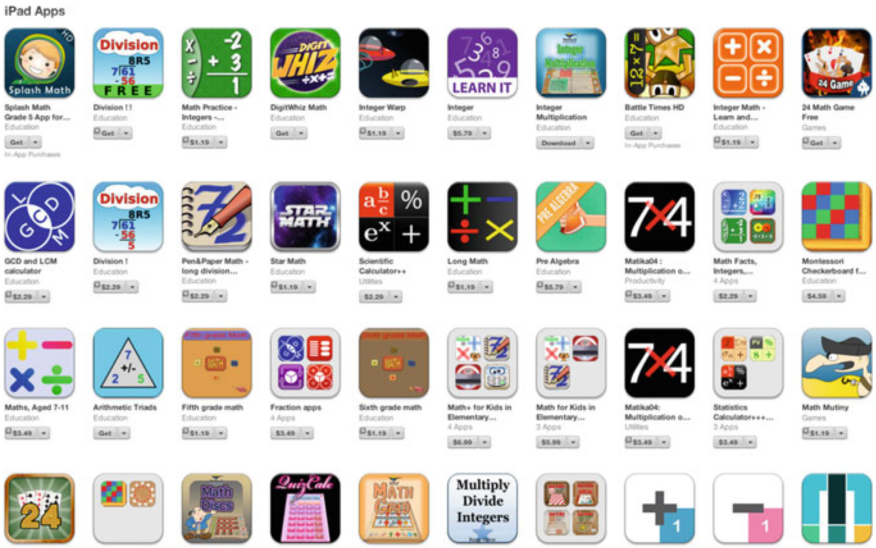


Fig. 12.3 iTunes store apps results for the keywords "integer multiplication"—iPad apps

relevant mathematics apps in other sections of the app store such as in apps for kids and edutainment. Because we were searching for apps for older children, we limited the scope of our search to the education section and to searching by key words.

We browsed all mathematics apps to select those that focused on learning negative integers as a curriculum area. We used the U.S. regional app store, although we also browsed the Canadian app store. For each of the apps in the results, from the keyword search and categories search, we examined the names and icons, as well as pulled up the iTunes App store pages of the app, to ascertain if the app fit the criteria of addressing negative integers. If an app showed a focus on positive integers, we also included it. The reason for this was because for some apps, the information available at the app store and at the app home page was not sufficient to show if an app on positive integers would extend to include negative integers. The home pages of the apps, where applicable, included more screenshots, detailed description of an app and, at times, video clips and reviews on an app. We eliminated all apps that did not focus on negative nor positive integers.

Selected searches by a keyword yielded a return of up to 100 results, the maximum possible, which pointed to the likelihood that more apps tagged with these keywords were available at the app store. To get a sense of how many more apps were left out by the app store results of the first 100 featured apps, we browsed a third-party website that offered analytics of apps at app stores—App Annie. App Annie returned 2024 iPhone apps and 1978 iPad apps for the search keyword "mathematics." It also returned 189 iPhone apps and 166 iPad apps for the keyword "integer." No apps were returned at App Annie when keywords were combined.

We selected 80 mathematics apps relevant to negative and positive integers (the Number Sense and Numeration strand) to download, try out, and review. The screenshots, descriptions, and information provided on an app were not always adequate for a review. We found that we had to download an app before we could ascertain its appropriate grade range and learning outcomes. Several app developers identified school grades, grade bands, or age groups for which the app was appropriate. Thirty-four of the 80 Number Sense and Numeration apps were found to be relevant to Grades 7 and 8; however, for many apps, the grades/ages indicated were not always accurate, at least not for the mathematics curriculum in the Canadian province where the research was conducted. Of the 34 apps that we found relevant to Grades 7 and 8 Number Sense and Numeration, only 8 were appropriately labelled as Grades 7, 8, or middle school apps. Overall, the grade bands indicated by the developers were not accurate. This is perhaps an indication that the developers are from varied countries where it is plausible that this content on negative numbers is addressed much earlier. Larkin (2013) interprets this as an indicator that the developers are not familiar with and do not consult a curriculum policy document when identifying grade fit of their app, or that the grade levels were selected from a marketing, rather than a curriculum, perspective. He found the targeted level to be 2–3 years younger than the ages specified by the app developer.

12.6 Dimensions for Selecting Appropriate Integer Apps

It was evident from the review of the 80 apps that several dimensions, including the nature of the curriculum addressed, were central when evaluating apps.

12.6.1 The Nature of the Curriculum Addressed

Emergent and adaptable activities as contrasted with overly prescriptive activities. We considered the nature of the learning that the mathematics tasks in the app could evoke. Only 3 of the apps involved what we refer to as, after Heydon and Wang (2006), emergent features (e.g., Math Alchemist Lite, and its other two versions). Math Alchemist is an example of an app that focused on a problem-solving context, the one of making 24, using any random numbers combined with number operations. A user's response becomes part of the inputs available for use in making 24, and the level of difficulty is increased depending on the user's success at a level. Apps with *emergent* features, ranked Level III apps on the curricula dimension, presented some rich mathematics problems that were, for instance, closely aligned with teaching through problem solving.

We labelled, again after Heydon and Wang (2006), *adaptable* apps as those 13 apps (e.g., Math Blaster Hyper Blast, Math Boosting, Interactive Integers) that posed questions or problems, which could have involved computing answers, but at least offered ways for the user to extend the problem. The Interactive Integer app posed tasks that involved conceptual understanding (see activities with coloured tiles) in addition to drill tasks for practicing integer addition and subtraction. This app was ranked Level II on the curricula dimension.

A majority of the apps, 74 out of 80, mainly offered prescriptive tasks. *Prescriptive apps* only posed traditional, prescriptive practice tasks, such as the question “3 + (−4) = ?” These focused on right/wrong responses from the learner in a manner similar to physical flash cards. These apps scored low, Level I, on the curricula dimension.

Only 4 apps (e.g., Interactive Integers, Math 24 Solver, and Math Blaster HyperBlast game) focused on building understanding of concepts, introducing a new topic, or explaining how a procedure worked, scoring high—Level III—on this dimension. A large number of apps, 58 out of 80, were for practicing earlier learned concepts, as would be the case with flash cards. That a majority of apps mainly offered prescriptive tasks was also the case in Larkin’s (2014) evaluation in which they found that procedural apps dominated.

Mathematics content aligned with more recent, higher quality curricula. Each of the 80 apps, according to their developers, was for learning, practicing, or getting quizzed on mathematical topics. The mathematics topics were listed differently, fluctuating from mentioning a single topic to listing a range of up to five topics. The topics included naming of general mathematics branches, such as arithmetic, through indicating a specific mathematics topic, such as negative numbers, to, at the highest ranking, Level III, further specifying mathematics content and learning outcomes (or, expectations), such as using models with negative integers. We view the latter focus that goes beyond naming a branch of mathematics or listing topics to specifying what is learned or practiced by using the app as a use of language consistent with that used in more contemporary, higher quality curricula of Canadian provinces and several other countries. In many curriculum documents, such as the NCTM principles and standards (NCTM 2000), the content specified goes beyond a mere mention of a topic to specifying learning expectations.

A selection of apps (e.g., Math 1st–6th Grade Digital Workbooks—Space Board) showed coverage for other strands, such as Geometry, in addition to Number Sense and Numeration. We took this focus, on connections of number sense to geometric representations of numbers, to align with the NCTM standards focus on connections among strands.

12.6.2 Actions and Interactions Afforded by the App

Constructive, Manipulable as Opposed to Largely Instructive Apps. Some adaptable apps involved interfaces with objects such as a number line that a user could act upon, or manipulate. In Figs. 12.4, 12.5, 12.6, 12.7, 12.8, 12.9 and 12.10, we show screenshots of the Interactive Integer app to illustrate how the number line and integer tiles in this app could be dragged and dropped as the user added or

Interactive Integers – Addition and Subtraction

Perfect for learning how to add and subtract positive and negative numbers!

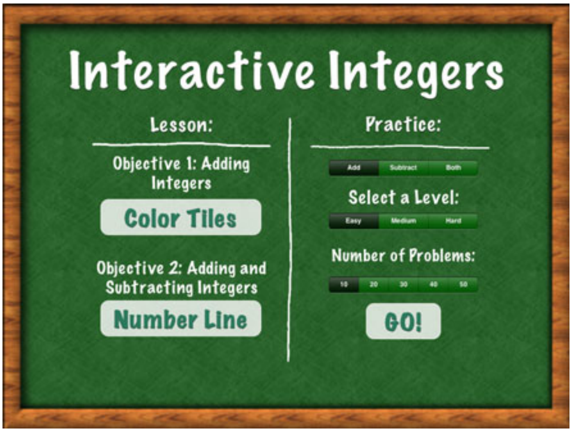
This app provides two fun interactive methods for learning and understanding addition and subtraction of positive and negative numbers. The color tile model and number line model are two of the most common classroom methods for teaching integers. These methods help students visually understand integer concepts instead of relying only on memorization of rules.

"The directions and explanations are very clear. I even let my second grader play around with this app, and

Fig. 12.4 Interactive Integers app—both iPhone and iPad app. *Source* www.tictaptech.net/apps/interactive-integers/

subtracted integers including negative integers. The colored tiles in the interactive integer apps could be dragged to demonstrate the identity property (e.g., +1 + −1 = 0): When a yellow, positive tile and a red, negative tile were dragged close to each other they each disappeared. Many representations of mathematics concepts in instructive apps could not be acted on or modified. Some apps only included audio or video demonstrations of an instructor explaining a mathematics process or giving the answer. A good number of apps did not have any objects that visually represented mathematics concepts. Goodwin and Highfield's (2013) evaluation found that a majority apps were instructive.

Fig. 12.5 Interactive Integers app showing user choice on task



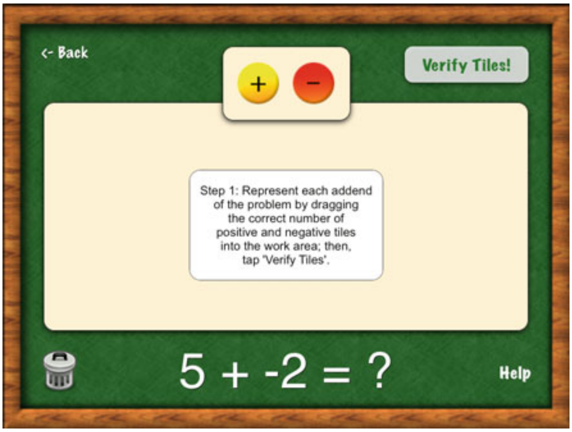


Fig. 12.6 Interactive Integers app color tile instructions

Step 1: Represent each addend of the problem by dragging the correct number of positive and negative tiles into the work area; then, tap 'Verify Tiles'.

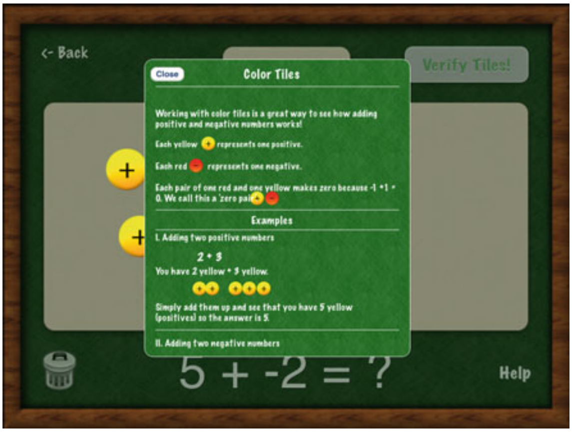


Fig. 12.7 Interactive Integers app hint on using color tiles

Color Tiles

Working with color tiles is a great way to see how adding positive and negative numbers works!

Each yellow ⊕ represents one positive.

Each red ⊖ represents one negative.

Each pair of one red and one yellow makes zero because -1 +1 = 0. We call this a 'zero pair' ⊕⊖

Examples

I. Adding two positive numbers

2 + 3

You have 2 yellow + 3 yellow.

⊕⊕ ⊕⊕⊕

Simply add them up and see that you have 5 yellow (positives) so the answer is 5.

II. Adding two negative numbers

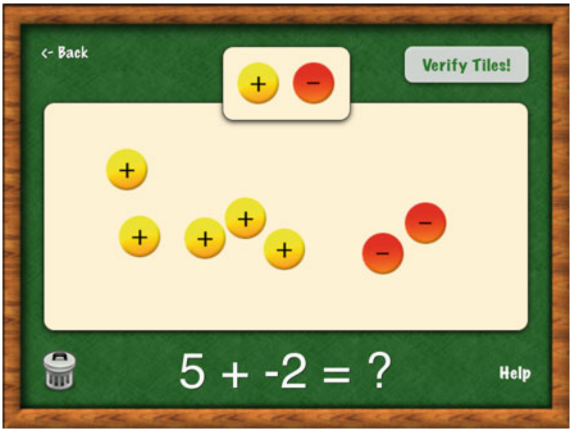
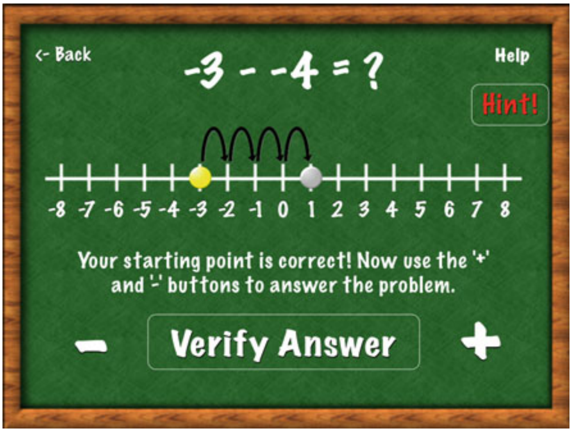
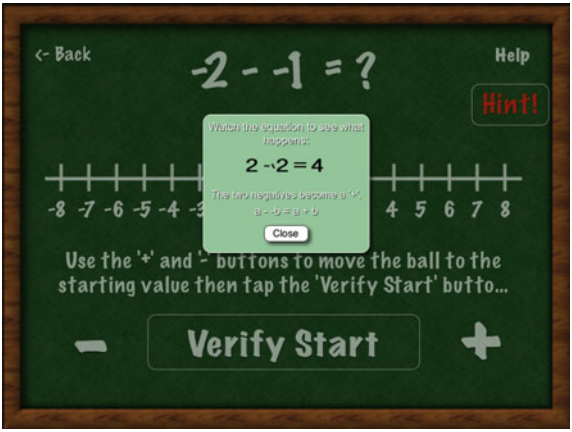


Fig. 12.8 Interactive Integers app adding 5 + −2 using dynamic counters



Your starting point is correct! Now use the '+' and '-' buttons to answer the problem.

Fig. 12.9 Interactive Integers app showing the number line model



Watch the equation to see what happens:

2 -‹2 = 4

The two negatives become a '+'.
a - -b = a + b

Use the '+' and '-' buttons to move the ball to the starting value then tap the 'Verify Start' butto...

Fig. 12.10 Interactive Integers app explaining a rule on taking away a negative integer

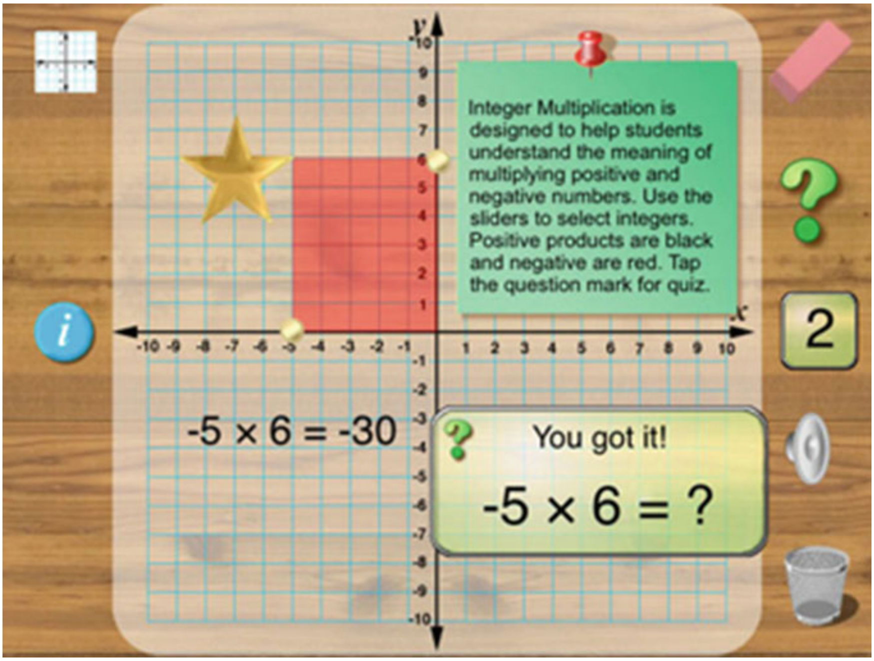
12.6.3 The Level of Interactivity and Range of Options Offered to the User

Multiple interaction apps as opposed to mono interaction apps. Only a few apps (e.g., Math Fact Master, Math!!!, and Middle School Math Pro 7th Grade) included opportunities for multiple users, such as submission of responses or marks, and asynchronous teacher interaction with the learner. We ranked apps with multiple interactions as high, Level III, on the dimension of interactivity and range of options, to be contrasted with apps offering mono interactions. Seventy of the 80 apps, including many of the apps that ranked Level III and Level II on the other dimensions, were designed with a focus on one user—mono interaction—thus limiting interaction to one user and the interface. In reference to video games, Gee (2012) distinguishes between the piece of software together with all the social activity around it and the piece of software alone. He would refer to the social activity around an app and the app as a software as the *Big A* app in contrast to the *small a* app because the former is important for participation, production and pro-active learning.

High- and moderate-engagement and interactivity apps as contrasted with low-engagement apps. Only 3 apps (e.g., Math 1 On-Track, Math Book Pro, and Math Blaster HyperBlast) accommodated a variety of inputs and choices, and offered varied possibilities of inputs and choices so the user may insert and select options, thus ranking Level III on interactivity. We referred to apps with a range and adjustability of options as high-interactivity apps. A good number of apps (e.g., Mathopolis, Math 2112, Math 24 Solver, Math4Touch), specifically 55 of 80, involved moderate interactivity with some opportunities for the users to input values and make choices. With the Interactive Integers app a learner was offered a choice of representation—tiles or the number line; the operation—addition or subtraction; number of questions; and level of difficulty. About a quarter, 22 of the 80 apps, involved much lower-interactivity, Level I. Many apps were limited to already inputted values and allowing only up to two choices (e.g., check answer and a "next" button) for the user.

12.6.4 The Quality of the Design Interface and Graphics in the App

Multi-media, high quality apps as opposed to primarily text-based, low quality apps. Sixty-six of the 80 apps utilized visual representations and graphics in addition to numeric symbols and text. Only 20 of the apps (e.g., Math Blaster HyperBlast, Interactive Integers, Integers, and Math!!!) went beyond using numerical symbols and text to utilize other mathematical representations such as geometric, graphic, simulations, or 3D graphs. We ranked these apps as Level III apps on technical design features. Forty-four of the 80 apps utilized sound effects and music. Many of these apps utilized multiple colors. Some apps used the colors in ways that were not simple add-ons, but in ways integral to the mathematics content. For instance, in the Integer Multiplication app, an iPad only app (see Fig. 12.11), the use of colors offered ways for the learner to identify patterns and distinguish characteristics of negative and positive integers. Still, a majority of apps, over 60 out of 80, largely utilized, at the lowest rank—Level I, only numerals and text to represent mathematics concepts. Haugland (2005) warns against this "poor use of a powerful learning tool" (p. 330).



Integer Multiplication is designed to help students understand the meaning of multiplying positive and negative numbers. Use the sliders to select integers. Positive products are black and negative are red. Tap the question mark for quiz.

-5 × 6 = -30

You got it!

-5 × 6 = ?

Fig. 12.11 Integer Multiplication app showing use of color

12.6.5 Instrument Content and Value Added by the Instrument

Virtual-only innovations, virtual developments of, with added value on, existing instruments as opposed to digitized images of existing materials. Because virtual and physical materials complement each other (Namukasa et al. 2009), for each of the apps, we examined the relation, if any, to existing instructional material/ resources. The team assessed if an app replicated already existing mathematics resources, such as virtual manipulatives, textbooks, or web resources, or whether an app was a digital version of these materials. This was important in assessing the app's pedagogical and cognitive elements (i.e., whether, for instance, it replicated a material that focused on developing conceptual knowledge, or on test preparation). Base-Ten Blocks replicated the physical and virtual Base-Ten Blocks manipulative. According to Bos (2009b), interactive mathematics learning tools, such as virtual manipulatives that are enhanced with technology, have a higher degree of cognitive fidelity than technology-based tools that focus on games, instructional information and quizzes. The representations of colored tiles and number lines, as seen in the Interactive Integers app, reflect the use of virtual, visual, and physical representations of integers in ways that are enhanced to represent a mathematics property. We found that many apps were designed based on mathematics puzzles (e.g., Math 24

Solver). Some apps added game contexts to paper-and-pencil mathematics puzzles. Also, many apps were game based (e.g., Mathopolis). Sixty-six apps involved some recreational features and 4 of these involved role-playing games (e.g., Math Blaster HyperBlast). Certain apps (e.g., YourTeacher, Motion Math-Zoom), in a manner similar to a mathematics textbook chapter or a lesson in a course, were part of a collection of apps focusing on varied mathematics topics for the same age level. Larkin (2013, 2015a, b) and Calder (2015) observed that many apps were stand-alone apps focusing on one particular kind of skill, knowledge, or content. Further, apps in bundles appeared to be aligned with curricula expectations.

12.6.6 The Level of Intellectual/Cognitive Involvement It Evokes—Intense, Limited or Very Low

Intense as opposed to limited or very low intellectual/cognitive involvement. Overall, apps with adaptable (or, emergent) characteristics and those with manipulable (or, modifiable) elements appeared to have the potential for intense intellectual/cognitive involvement whereas apps with instructive and prescriptive characteristics appeared to have limited to very low potential for intense intellectual/cognitive involvement. Even among prescriptive and instructive apps, some apps, because they scored high on other dimensions such as on interactivity and range of options and technical design aspects, appeared to be more engaging and thus offered potential for intellectual involvement at the procedural level.

A good number of apps combined elements on one dimension as illustrated in Goodwin and Highfield (2013) and Gadanidis et al. (2004). We did not find an app that ranked at level III for all dimensions. The Interactive Integers app combined both the adaptive and prescriptive elements on the curricular dimension, and it had a manipulable interface (level II on the actions and interactions dimension). It also offered choice and provided immediate feedback, as well as written instruction for both the lessons on understanding and for practice questions, but did not offer an opportunity for the learner to input values or make annotations by including a keypad. One of its instructions on how to take away a negative number was not mathematically accurate. Interactive Integers was limited to integer subtraction and addition. The Integer Multiplication app, that scored high on the characteristic of use of color to focus learning, covered only a single operation on integers—multiplication.

Some apps that scored low, Level I, on one dimension scored higher, Level II or III, on other dimensions. Even when it focused on right and wrong answers—Mathopolis, a prescriptive app—also involved a game context that allowed user choice on the level of difficulty and nature of operations, scoring Level II on interactivity. One could say that Mathopolis scores high among prescriptive apps because it is a Level II app on at least one other dimension. One of the apps that appeared to involve emergent features had a game context that did not appear appropriate for middle school students. We pondered the messaging and content in the apps and its appropriateness for learners. This was also the case in Larkin (2015a), where he found apps scoring high on one dimension and low on another. For instance the apps Larkin evaluated scored higher on pedagogical fidelity, followed by mathematical fidelity, and lowest on cognitive fidelity. To Haugland (1999), children's software should be evaluated on age appropriateness and non-violence.

Apps with multiple interactions—between several users (e.g., Math Fact Master which could submit scores to an email address, as well as Math!!! with the possibility of a teacher embedding messages) have promising added value of interacting with others through the cognitive tool.

12.7 Concluding Remarks

Our evaluation instrument could guide teachers when selecting apps that meet their teaching goals. As well, the evaluation instrument could guide developers in designing apps that are more aligned with emergent and adaptive curriculum, that also focus on conceptual understanding in addition to focusing on procedural and declarative knowledge, and that utilize multiple and interactive modes in ways that are central to the representation of mathematics entities.

Some teachers implement and test objects, many use objects recommended by colleagues, and yet other teachers, especially those comfortable with computer programming, increasingly approach the use of learning objects from a developer's perspective. New friendly coding programs are making it easier for more teachers, and even students, to engage in designing apps. Thus, our instrument can potentially guide students, teachers, educators, and researchers when they design apps.

When mathematics apps are thoughtfully used in ways that encourage learners to do the mathematics (i.e., explore, conjecture, test, and apply), rather than only doing procedural steps, learning apps have the potential to deepen mathematical understanding and encourage students to work at higher levels of generalization and abstraction (Bos 2009a, b). Looking to the future, with the increased focus on students of all ages learning to code, such as the mandate of coding across all grades in England's National Curriculum (UK Government News Release, 4 February 2014), we need to also consider: (1) the connection between students as coders and students as mathematics learners, and (2) the design of apps, not only as education products to be consumed, but also environments that may be edited and reprogrammed by users. For example, Gadanidis and Yiu (2014) created HTML5 apps (available at www.researchideas.ca/mathncode) that attempt to meet these conditions, respectively, by: (a) using app interfaces where users change code parameters to control a simulation or play a game, and (b) programming apps in MIT's Scratch environment, giving students full access to the code, which they can edit to create variations or new simulations and games. Explicitly incorporating coding in mathematics apps would help incorporate three pedagogical benefits of coding in mathematics learning: making concepts tangible, making relationships dynamic, and giving students more control over the learning process (Gadanidis 2014, 2015).

References

Bartolini, M. G., & Martignone, F. (2014). Manipulatives in mathematics education. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 365–372). Dordrecht, Netherlands: Springer.

Basham, J., Meyer, H., & Perry, E. (2010). The design and application of the digital backpack. *Journal of Research on Technology in Education, 42*, 339–359.

Bortolossi, H. J. (2012). Criando conteúdos educacionais digitais em matemática e estatística com o uso integrado de tecnologias GeoGebra, JavaView, HTML, CSS, MathMLe JavaScript. *Revista do Instituto GeoGebra de São Paulo, Journal od the Sao Paulo GeoGebra Institute, 1* (1), 38–39.

Bos, B. (2009a). Technology with cognitive and mathematical fidelity: What it means for the math classroom. *Computers in the Schools, 26*(2), 107–114.

Bos, B. (2009b). Virtual math objects with pedagogical, mathematical, and cognitive fidelity. *Computers in Human Behavior, 25*, 521–528.

Bos, B. (2011). Professional development for elementary teachers using TPACK. *Contemporary Issues in Technology and Teacher Education, 11*(2), 167–183.

Calder, N. (2015, October). Apps: Appropriate, applicable, and appealing? In T. Lowrie & R. Jorgensen (Eds.), *Digital games and mathematics learning* (pp. 233–250). Dordrecht, Netherlands: Springer.

Cayton-Hodges, G., Feng, G., & Pan, X. (2015). Tablet-based math assessment: What can we learn from math apps? *Journal of Educational Technology & Society, 18*(2), 3–20.

Dick, T. P. (2008). Keeping the faith: Fidelity in technological tools for mathematics Education. In G. W. Blume & M. K. Heid (Eds.), *Research on technology and the teaching and learning of mathematics: Vol. 2. Cases and perspectives.* (pp. 333–339). Charlotte, NC: Information Age.

Dickens, H., & Churches, A. (2012). *Apps for learning: 40 best iPad/iPod Touch/iPhone apps for high school classrooms*. Vancouver, BC: 21st Century Fluency Project.

Fullan, M., & Donnelly, K. (2015). *Evaluating and assessing tools in the digital swamp*. Bloomington, IN: Solution Tree Press.

Gadanidis, G. (2014). Young children, mathematics and coding: A low floor, high ceiling, wide walls learning environment. In D. Polly (Ed.), *Cases on technology integration in mathematics education* (pp. 312–344). Hersey, PA: IGI Global.

Gadanidis, G. (2015). Coding as a Trojan horse for mathematics education reform. *Journal of Computers in Mathematics and Science Teaching, 34*(2), 155–173.

Gadanidis, G., & Schindler, K. (2006). Learning objects and embedded pedagogical models. *Computers in the Schools, 23*, 19–32.

Gadanidis, G., Sedig, K., & Liang, H. N. (2004). Designing online mathematical investigation. *Journal of Computers in Mathematics and Science Teaching, 23*(3), 273–296.

Gadanidis, G., & Yiu, C. (2014). *Math and code*. Retrieved from www.researchideas.ca/mathncode

Gee, J. P. (2005). Good video games and good learning. *Phi Kappa Phi Forum, 85*(2), 33–37.

Gee, J. P. (2012). Digital games and libraries. *Knowledge Quest, 41*(1), 61–64.

Goodwin, K., & Highfield, K. (2012). *iTouch and iLearn: An examination of "educational" apps*. Paper presented at Early Education and Technology for Children Conference, March 14–16, 2012, Salt Lake City, Utah, USA.

Goodwin, K., & Highfield, K. (2013). A framework for examining technologies and early mathematics learning. In L. D. English & J. T. Mulligan (Eds.), *Reconceptualizing early mathematics learning* (pp. 205–226). New York, NY: Springer.

Haugland, S. W. (1999). The newest software that meets the developmental needs of young children. *Early Childhood Education Journal, 26*(4), 245–254.

Haugland, S. W. (2005). Selecting or upgrading software and websites in the classroom. *Early Childhood Education Journal, 32*(5), 329–340.

Heydon, R., & Wang, P. (2006). Curricular ethics in early childhood education programming: A challenge to the Ontario kindergarten program. *McGill Journal of Education, 41*(1), 29–46.

Highfield, K., & Goodwin, K. (2013). Apps for mathematics learning: A review of 'educational' apps from the iTunes App Store. In V. Steinle, L. Ball, & C. Bardini (Eds.), *Mathematics education: Yesterday, today and tomorrow*. Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia. Melbourne, VIC: MERGA.

Hitt, F. (Ed.). (2002). *Representations and mathematics visualization*. Mexico: PME-NA, Cinvestav-IPN.

Kamii, C., Lewis, B. A., & Kirkland, L. D. (2001). Fluency in subtraction compared with addition. *Journal of Mathematical Behavior, 20*(1), 33–42.

Kay, R. H., & Knaack, L. (2009). Assessing learning, quality and engagement in learning objects: The learning object evaluation scale for students (LOES-S). *Educational Technology Research and Development, 57*(2), 147–168.

Larkin, K. (2013). Mathematics education. Is there an app for that? In V. Steinle, L. Ball, & C. Bardini (Eds.), *Mathematics education: Yesterday, today and tomorrow.* Proceedings of the 36th annual conference of the Mathematics Education Research Group of Australasia (pp. 426–433). Melbourne, VIC: MERGA.

Larkin, K. (2014). iPad apps that promote mathematical knowledge? Yes, they exist! *Australian Primary Mathematics Classroom, 19*(2), 28–32.

Larkin, K. (2015a). The search for fidelity in geometry apps: An exercise in futility? In M. Marshman, V. Geiger, & A. Bennison (Eds.), *Mathematics education in the margins.* Proceedings of the 38th annual conference of the Mathematics Education Research Group of Australasia. Sunshine Coast, QLD: MERGA.

Larkin, K. (2015b). An app! An app! My kingdom for an app: An 18-month quest to determine whether apps support mathematical knowledge building. In *Digital games and mathematics learning* (pp. 251–276). Dordrecht, Netherlands: Springer.

Maddux, C., Johnson, D., & Willis, J. (2001). *Educational computing: Learning with tomorrow's technologies*. Needham Heights, MA: Allyn and Bacon.

Mills, M., Goos, M., Keddie, A., Honan, E., Prendergast, D., Gilbert, R., & Renshaw, P. (2009). Productive pedagogies: A redefined methodology for analyzing quality teacher practice. *Australian Educational Researcher, 36*(3), 67–87.

Moyer, P. S., Bolyard, J. J., & Spikell, M. A. (2002). What are virtual manipulatives? *Teaching Children Mathematics, 8*(6), 372–377.

Moyer-Packenham, P. S., Salkind, G., & Bolyard, J. J. (2008). Virtual manipulatives used by K–8 teachers for mathematics instruction: Considering mathematical, cognitive, and pedagogical fidelity. *Contemporary Issues in Technology and Teacher Education, 8*(3), 202–218.

Moyer-Packenham, P. S., Shumway, J. F., Bullock, E., Tucker, S. I., Anderson-Pence, K., Westenskow, A., et al. (2015). Young children's learning performance and efficiency when using virtual manipulative mathematics iPad apps. *The Journal of Computers in Mathematics and Science Teaching, 34*(1), 41–69.

Moyer-Packenham, P. S., & Westenskow, A. (2013). Effects of virtual manipulatives on student achievement and mathematics learning. *International Journal of Virtual and Personal Learning Environments, 4*(3), 35–47.

Namukasa, I. K., Stanley, D., & Tutchie, M. (2009). Virtual manipulative materials in secondary mathematics: A theoretical discussion. *Journal of Computers in Mathematics and Science Teaching, 28*, 277–307.

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

Pelton, L. F., & Pelton, T. (2012, March). Sharing strategies with teachers: iPods in math class. In *Society for information technology & teacher education international conference* (Vol. 2012, No. 1, pp. 4363–4366).

Pepin, B., & Gueudet, G. (2014). Curriculum resources and textbooks in mathematics. In S. Lerman (Ed.), *Encyclopedia of mathematics education* (pp. 132–135). Dordrecht, Netherlands: Springer.

Peterson, J. C. (1972, May). Fourteen different strategies of multiplication of integers or why ($^{-}$1) ($^{-}$1)=$^{+}$1. *The Arithmetic Teacher*, *19*(5), 397–403.

Pirie, S., & Kieren, T. (1994). Growth in mathematical understanding: How can we characterize it and how can we represent it? *Educational Studies in Mathematics, 26*(2–3), 165–190.

Sedig, K., Parsons, P., Dittmer, M., & Haworth, R. (2014). Human-centred interactivity of visualization tools: Micro- and macro-level considerations. In W. Huang (Ed.), *Handbook on human centric visualization* (pp. 717–743). New York, NY: Springer.

Skip, J. (1990). But everybody accepts this explanation: Operations on signed numbers. In J. Fauvel (Ed.), *History in the mathematics classroom (The IREM papers)* (Vol. 1). London, England: Mathematical Association.

Trouche, L., Drijvers, P., Gueudet, G., & Sacristan, A. I. (2013). Technology-driven development and policy implications for mathematics education. In A. J. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick, & F. K. S. Leung (Eds.), *Third international handbook of mathematics education* (pp. 753–790). New York, NY: Springer.

UK Government News Release. (February 4, 2014). Year of code and £500,000 fund to inspire future tech experts launched. Retrieved from www.gov.uk/government/news/year-of-code-and-500000-fund-to-inspire-future-tech-experts-launched

Yerushalmy, M., & Ben-Zaken, O. (2004). Mobile phones in education: The case for mathematics, Haifa. Retrieved from http://construct.haifa.ac.il/~michalyr/celular%20report.pdf

Zbiek, R. M., Heid, M. K., Blume, G. W., & Dick, T. P. (2007). Research on technology in mathematics education: A perspective of constructs. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 2, pp. 1169–1207). Charlotte, NC: Information Age.

Zhang, M., Trussell, R. P., Gallegos, B., & Asam, R. R. (2015). Using math apps for improving student learning: An exploratory study in an inclusive fourth grade classroom. *TechTrends, 59* (2), 32–39.