The Search for Fidelity in Geometry Apps: An Exercise in Futility?

Kevin Larkin
Griffith University
<k.larkin@griffith.edu.au>

As the use of mathematics apps in classrooms becomes more prevalent, robust research into their effectiveness is required to inform best practice regarding their use. This is particularly the case for Geometry apps where accurate and dynamic representations are critical in enhancing mathematical learning. This paper provides findings from an initial critique of 53 Geometry apps. Early findings indicate that the majority of these apps were limited in their ability to assist students in developing Geometrical conceptual understanding; however, all is not lost as a number of apps were highly appropriate.

This paper briefly synthesises the research literature concerning the use of digital manipulatives and then outlines the qualitative component of a broader three-step methodology for critiquing the appropriateness of Geometry apps. Early findings of the research indicate that the majority of the iPad Geometry apps reviewed would do little to assist students in developing Geometry concepts. Research into apps is needed as, although there has been some research into the mathematical effectiveness of apps (See Attard & Curry, 2012; Larkin, 2013) there has been little to no specific research into their usefulness in developing Geometry concepts. In addition, much of the current research into apps, with some exceptions (Larkin, 2014, 2015; Moyer-Packenham, et al., 2015) has been largely descriptive. An initial review of apps (Larkin, 2013) uncovered very few Geometry specific apps; however, the app market has since burgeoned with the creation of a range of Geometry specific apps. For the purpose of this paper, Geometry apps are those that include content from the Geometry sub-strands of the Australian Mathematics Curriculum. As was indicated in Larkin (2014), determining the quality of an app is difficult not only because of the lack of research, but also because the information that is available at the iTunes Appstore is written by the app developers to sell their app and thus not reliable. The problem of determining app quality in relation to Geometry is additionally complex as these apps require the creation of mathematically accurate external representations. Earlier research (Larkin, 2013) suggests that accuracy in representations was not commonly evident; consequently, a new methodology for evaluating Geometry apps was designed. This paper outlines how the constructs of pedagogical, mathematical and cognitive fidelity were used to evaluate 53 Geometry specific apps. The goals of this paper are two-fold. Firstly, to articulate a component of a broader methodology for reviewing the apps such that other researchers can use the methodology; and secondly, as reviewing apps is a time consuming process, an outcome of the research was the creation of a web-based database, available to teachers, of Geometry apps. This research recognises that the selection of appropriate Geometry apps needs to be based on a deeper understanding of the strengths and weaknesses of the apps, and what makes them pedagogically, mathematically and cognitively reliable (Bos, 2009).

Literature Review

It is taken as given in this paper that manipulatives (concrete and digital) support mathematical learning (e.g., Burns & Hamm, 2011; Carbonneau, Marley, & Selig, 2013; Moyer-Packenham, et al., 2015; Özel, Özel, & Cifuentes, 2014). This affords space to

2015. 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), pp. 341–348. Sunshine Coast: MERGA.

address more fully the research on three aspects of fidelity required in Geometry apps; namely, pedagogical, mathematical, and cognitive fidelity (Dick, 2008).

Pedagogical, Mathematical and Cognitive Fidelity

Pedagogical fidelity is defined by Dick (2008) as the degree to which a student can use a tool to further their learning. Zbiek, Heid, Blume and Dick, (2007) suggest that pedagogical fidelity also refers to "the extent to which teachers (as well as students) believe that a tool allows students to act mathematically in ways that correspond to the nature of mathematical learning that underlies a teacher's practice" (p.1187). Dick (2008) suggests that a pedagogically faithful tool will likely be described by students in terms of how it allowed them to interact with mathematics (e.g., "I created this triangle" etc.) rather than simply as a description of procedures for use (e.g. "I set the preferences to the fastest level" etc.). Therefore, to be an effective pedagogical tool, an app must support any action by the student that will lead to conceptual understanding of the underpinning mathematical principle.

The second aspect of fidelity to consider is mathematical fidelity. Zbiek et al. (2007) defines it as the "faithfulness of the tool in reflecting the mathematical properties, conventions, and behaviors (as would be understood or expected by the mathematical community)" (p.1173). Thus, mathematical fidelity is present when the activity of a student "is believable, is concrete, and relates to how mathematics is a functional part of life" (Bos, 2011, p. 171); and when they add strength to an understanding of mathematics as a language of patterns and order. Dick (2008) cautions that the drive for user friendliness can sometimes run contrary to faithfulness to an accurate mathematical structure. This is particularly worrisome as most apps are designed for (a) market reasons and (b) by non-educators (Larkin, 2013). Keeping the notion of mathematical fidelity at the forefront of decisions when selecting apps reminds teachers to avoid apps that do not deliver accuracy in terms of mathematical content or constructs e.g. correct scaling may not be evident in transformations.

The final aspect of fidelity is cognitive fidelity, which refers to "the faithfulness of the tool in reflecting the learner's thought processes or strategic choices while engaged in mathematical activity" (Zbiek et al., 2007, p.1173). Cognitive fidelity can be viewed largely in terms of the external representations provided by the tool. Zbiek et al. further note that "if the external representations afforded by a cognitive tool are meant to provide a glimpse into the mental representations of the learner then the cognitive fidelity of the tool reflects the faithfulness of the match between the two" (p.1176). This notion of cognitive fidelity is critical in Geometry apps which are likely to utilise many external representations. The digital nature of "app objects" potentially results in high levels of cognitive fidelity, for example, 3D objects can be pulled apart and put back to together, and in so doing, reinforce the link between 3D objects and their 2D representations (i.e. nets); however, we will see that such potential is often unrealised in Geometry apps.

Although an understanding of the three types of fidelity can assist teachers in making decisions about whether or not to use an app, I have argued above that an issue for teachers is the time required to determine app quality via the three fidelities or other evaluative measures. In addition, although it might be expected that many of the findings on the use of virtual manipulatives would reflect the experience of using mathematics apps, rigourous quantitative research into mathematics apps is still in its infancy (Larkin, 2015; Moyer-Packenham et al., 2015). Therefore research into Geometry apps, which might be best placed to take advantage of the iPad's representational capability, is required.

Methodology

This section outlines the process for initially finding the Geometry apps and then explains how a qualitative review and a descriptive, quantitative measure of fidelity were used to evaluate the apps.

Locating and Scoring the Apps

Evaluation of the apps commenced with a targeted search for mathematics apps at the iTunes Appstore in October, 2014. The following search terms were used: Geometry Elementary Education; Geometry Junior Education; Geometry Primary Education; Symmetry Education; and Transformations Education. Many of the apps appeared in two or more of the searches.

Table 1
Levels of Fidelity in Geometry Apps – Adapted from (Bos, 2009)

Type of Fidelity	**Low Level (1-3)**	**Medium Level (4-7)**	**High Level (8-10)**
Pedagogical (Including Technological) The degree to which the App can be used to further student learning.	App is difficult to work with. Accessing all aspects of the app is difficult. App is not appropriate for the mathematics concepts it uses. Transitions are inconsistent or illogical.	Using App is not initially intuitive; but with practice becomes so. Mathematical activities presented are appropriate but could be developed without app. Transitions evident but only made via trial & error.	Manipulation of App is intuitive & encourages user participation. Little or no training or instructions are required. Transitions are logical & aid sense making.
Mathematical The degree to which the App reflects mathematical properties, conventions and behaviours.	Mathematical concepts are underdeveloped or overly complex. Lack of patterns. Lack of connection to real world mathematics.	Application of mathematics concepts unclear. Patterning is evident but lacks predictability or is unclear. Some connection to real world mathematics.	Mathematics concepts developed are correct & age appropriate. Patterns are accurate & predictable. Clear connection with real world mathematics.
Cognitive The degree to which the App assists the learner's thought processes while engaged in mathematical activity.	No opportunities to explore or test conjectures. Static or inaccurate representations. Patterns do not connect with concept development.	Limited opportunities to explore or test conjectures. Minor errors with representations but still make sense. Patterns connect in a limited way with concept development.	App encourages exploration & testing of conjectures. Representations are accurate & easily manipulated. Patterns clearly aide concept development.

Apps were excluded from the final review according to a variety of criteria whereby only one app in any series was reviewed and apps categorised as Games, Entertainment or Lifestyle; apps where mathematics was part of a larger package of reading, writing, and spelling skills; and apps that required additional costs for access or further online registration were excluded.

As indicated earlier, Dick's (2008) three dimensions of pedagogical, mathematical, and cognitive fidelity have been used by other researchers to determine the quality of mathematics manipulatives (e.g., Bos, 2009; Zbiek et al., 2008). Two methodological innovations in this research are using the measures to evaluate apps; and the use of numerical values to represent the degree to which these three dimensions are present. Bos (2009) went some way towards using the dimensions as an assessment tool categorising software as Low, Medium and High fidelity in each dimension. Table 1 is an adapted version of Bos' work, modified specifically for evaluating Geometry apps. In order to make sophisticated comparisons between the three dimensions of fidelity, the nominal levels of Low, Medium and High have been replaced by an ordinal continuum ranging from 1 (no fidelity) to 10 (very high fidelity) for each of the three dimensions.

An app is considered low level (1-3) if it is generally static and inaccurate mathematically and fails to develop mathematical concepts. It is considered medium level (4-7) if more than one solution is possible and conjectures are possible (but not testable) and transitions between different aspects of the app are possible but unclear. Finally, an app is considered high level (8-10) if it uses accurate representations that are easy to manipulate with transitions between app elements that are logical and consistent, and it affords the formation of multiple, testable conjectures. In this evaluative schema, an individual app could score, for instance, highly on mathematical fidelity, yet poorly on cognitive or pedagogical fidelity.

Findings and Discussion

Prior to a brief discussion on the initial descriptive statistics collected in this research, a comprehensive qualitative evaluation (see Table 2) of the apps is provided. The author's prior research into the use of apps has indicated that this type of qualitative information is very important for teachers in making decisions about whether or not to use an app. The qualitative reviews of each of the 53 apps are available at (*link removed for peer review*). I have included below an example of one of the reviews.

Table 2
Example Qualitative Geometry App Review

App Name	**Content**	**Yr. Level**	**Generic Features of the App**
3D Geometry Basica	Shapes	Years 6-7	This app includes eight common 3D objects. The only action which can be performed on the objects is a simple zoom in or out. Each object includes a mathematical description in mathematics language and includes formulas for Surface Area and Volume.
Reviewer Comments re Mathematical Fidelity: Using the app is intuitive, largely due to the limited options available, and the content is accurate. From a conceptual development perspective the app contains complex formulas for finding SA and Volume in Platonic Solids, spheres and cylinders, but no linkage is established between the SA of an object and its Volume, or between the SA and Volumes of the different objects – E.G. between Pyramid and Octahedron. No connection to the real world. Static representations, no nets, no option for patterning or for testing conjectures. Very limited usefulness and the app does nothing that the actual physical objects couldn't do.			

As outlined in Larkin (2013) initially locating potential useful apps is a complex and time consuming process and therefore the provision of this qualitative review of each app is very useful for teachers. Apps are difficult to find due to the sheer number of apps [approx. 150 000 education apps at the iTunes store (148AppsBiz, 2015)] and this difficulty is compounded by mismatches with naming (name of app at iTunes store is different to name of app on iPad), similar naming (a dozen apps had variations on the word geometry), the rapid turnover of apps at the store, and finally a very poor search engine (apps not sorted according to category or alphabetically). As indicated in the 2013 research, teachers are extremely time poor and thus are likely, if they decide to use apps at all, to be guided by the description at the iTunes store. These are at best "infomercials" and provide misleading details about the app. For all these reasons, educationally robust reviews such as the one provided here are critical if teachers are to be directed to find what amounts to a "needle in a haystack" – i.e., an app that is appropriate for them to use with their students.

Provided in the following paragraphs are findings based on initial descriptive analysis of the data regarding types of app content, levels of quality according to each of the three fidelities, an analysis of the range of scores across the three fidelities, and finally a brief description of seven apps which scored above 6/10 for each fidelities indicating a high level of appropriateness for classroom use. Turning to content analysis first, Table 3 indicates the number of apps that included a range of Australian Curriculum Geometry content.

Table 3
Number of Apps Providing Australian Curriculum Geometry Content #

Sub-Strand / Concepts	No. of Apps	Sub-Strand | Concepts	No. of Apps
Lines (1D)	16*	Slide (Translate)	10
Shapes (2D)	31	Flip (Reflect)	21
Objects (3D)	17	Turn (Rotate)	16
Angles	15	Dilations	6

***NB**: Total app count exceeds 53 as a number of apps include more than one type of content and are therefore counted more than once. # Pythagoras and trigonometry is only introduced in Australian secondary schools and so was beyond the scope of this review.

A number of apps just focussed on one content area (e.g., Simitri – line symmetry); however, many others took a broad brush stroke approach and covered content from two or more areas (e.g. EZ Geometry or Jungle Geometry). This is not always an advantage as broad coverage often meant shallow conceptual development and less usefulness as only one section of the app was appropriate for a particular year level. By far the most popular content area was Shapes and this may be because many of the apps were targeted at very young students (Foundation and Early Years) and also because these apps appear easy to create from a technical perspective. Whilst most common, many of these Shapes apps were very basic and only included naming of the shapes and very simple matching exercises. Many of these activities could more easily be completed using actual shapes. Reflections were the most common of the four major transformations presented in the apps and this may be a consequence of the desire to link the apps to symmetry in nature or the built environment which is more easily represented than rotational symmetry, translations or dilations. Angles and 1D Geometry apps appear were common; however, this is a result of a large number of quiz apps (largely concerning geometric reasoning) rather than the availability of a large number of apps developing understanding of 1D and Angles.

Table 4 provides a breakdown of the number of apps scoring six or more in each of the three respective fidelities. Although this looks like a healthy number of apps (42) scoring at least one six, this is not the case as many of the better apps scored a six or more in two or three categories. Overall, 26 of the 53 apps failed to score a six in any category; the average score of the 53 apps was 12.9/30; and none of the three fidelity categories scored an average of 50%. This is a clear indication that there are a large number of Geometry apps, categorised as educational at the iTunes store, which do not even meet a very low benchmark for appropriateness in classrooms. As might have been anticipated [given the findings of previous research (Larkin, 2014; 2015) which indicated that many apps are instructional and focus on declarative or procedural knowledge], the apps which were of some use tended to score well on the pedagogical fidelity dimension, less well in terms of the quality of the mathematics they contain, and generally poorly in their ability to assist cognitive development. This again mirrors the generally poor level of conceptual knowledge developed by apps in the research noted above.

Table 4

Number of Apps Scoring 6 or More in Respective Fidelities

Type of Fidelity	**Number of Apps (n=53)**	**Percentage* (to nearest 0.1)**	**Average Score / 10**
Pedagogical	21	39.6%	4.9
Mathematical	13	24.5%	4.3
Cognitive	8	15.1%	3.7
Overall Average Score for Apps on the three measures / 30			12.9

Overall, the apps scored more highly in terms of pedagogical fidelity because this is the easiest of the categories for non-mathematical app designers to mimic in their apps. Many of the apps met one of the pedagogical criteria, namely, they were easy to use without instruction, and many of them partially met the criteria of appropriateness of activity without necessarily doing anything more than could be easily replicated with an IWB, physical manipulatives, or even pen and paper. Many of them incorporated multiple choice quizzes (of varying degrees of quality) which may serve some use as revision exercises. This was particularly the case where quizzes drew from a large bank of questions, did not allow multiple guesses, and allowed results to be emailed (e.g. Kids Math-Angle Geometry and Symmetry School Learning).

Mathematical fidelity issues generally related to incorrect naming or classification of shapes and objects, (e.g. diamonds instead of rhombuses, cubes not considered prisms, squares not considered as rectangles, triangles not included as polygons); use of prototypical shapes and standard orientations (only three apps focused on non-prototypical shapes – Cyberchase Quest, Maths Geometry, and Shapes MyBlee); and lack of connection to any notion of real world application of mathematics (minor exceptions to this include Geometry 4 Kids and Simitri).

Of most concern was the low cognitive fidelity of most apps and this is problematic in terms of classroom use as this relegates many of the apps to only being useful as revision activities of for rote learning. The majority of apps did not meet the criteria for supporting cognitive development. Despite being technically capable, most apps only provided static representations and, where dynamic representations were used, they did not mimic the physical activity of, for instance, turning or sliding or flipping but used arrows or numbers to direct the transformations (noteworthy exceptions were Squares and Colors and Shapes MyBlee). In addition, very few apps allowed opportunity for students to create patterns and develop their own conjectures regarding shapes, objects, angles, or transformations. This is a serious shortcoming of the vast majority of the apps.

Despite the comments above, it is not all doom and gloom in “AppLand” as there are some apps that shine in the overall geometric darkness that is the iTunes store (see Table 5). Of the apps reviewed, seven of them (13% of the total apps reviewed) scored six or more out of 10 for each of the three fidelities. These are clearly the apps that teachers should be utilising in their classroom practice. What is interesting here is that apart from the top three, even the better apps were inconsistent in meeting the three fidelity standards as four of the seven scored one six and two of these four scored two sixes.

Table 5
Apps that Scored 6 or More on Each of the Three Fidelities

App Name	**Pedagogical**	**Mathematical**	**Cognitive**	**Total**
Co-ordinate Geometry	9	8	9	26
Transformations	9	8	9	26
Attribute Blocks	8	8	8	24
Shapes – 3D Geometry	9	6	8	23
Shapes and Colors	7	6	7	20
Pattern Shapes	8	6	6	20
Isometry Manipulative	7	6	6	19

This level of inconsistency mirrors the findings of Moyer-Packenham et al. (2015) in relation to virtual manipulatives. In their research they noted multiple affordances within each virtual manipulative such that one or more of these affordances may be more influential and beneficial for student learning. An example of this in terms of apps is Isometry Manipulative, where one component of the apps is extremely beneficial whilst the second component, if used, is likely to undermine student learning. This inconsistency becomes more apparent as scores further down the total list of scores are examined, for example, Geometry Montessori (9, 6, 5) scored equal to or higher than three of the apps listed in the top seven but was relatively poor in terms of cognitive development. Three other apps scored highly in pedagogical and mathematical fidelity but poorly in terms of cognitive development (GeoEng- 8, 6, 5; Geometry 4 Kids- 8, 6, 3; and Geometry Explore- 6, 6, 4). It is worth noting that only one app (Simitri- 4, 9, 8) scored very lowly in pedagogical fidelity but very highly in mathematics and cognitive fidelity. This indicates that this app should not be used unsupervised by students; however, with correct scaffolding from the teacher, it is very useful for developing mathematical understanding due to its high level mathematical and cognitive fidelity.

It is clearly the case that, other than with the top three apps, teachers need to decide the exact purpose they want to achieve by using an app and then look at the content covered and individual fidelity scores of each app, to find one that meets that specific purpose. In this manner, Geometry Montessori would be most appropriate to use in a revision mode but less so in terms of developing mathematical or conceptual fidelity. The full list of scores is available for teachers at the URL provided earlier in the paper.

Limitations and Next Steps

As was the case in Larkin (2014), a limitation of any study reviewing apps is a consequence of two factors; initially locating (and relocating apps), and the nature of the iTunes App store. Firstly, the sheer number and method of labelling apps (e.g., multiple apps called Geometry [or very similar] or apps containing geometry but not indicated in their name – e.g., Koala Math) means that there may be useful Geometry apps not reviewed. Secondly, the iTunes store is a moveable feast as apps are generated, renamed, relocated, or removed on a daily basis. This research has indicated that, although many Geometry apps are quite poor in terms of their fidelity, it is, to return to the question posed in the title, certainly not a futile exercise to use some of them in primary mathematics classrooms. Due to the shortened nature of MERGA conference papers, only one component of the quantitative measures used in the broader research has been presented to support this claim. A more substantive examination of their quality incorporating three quantitative measures, using modified versions of Haugland's (1999) Software Scale, Bos' (2009) software categorisations and Dick's (2008) three fidelities will be used in future use to more comprehensively determine the quality of Geometry apps in supporting primary students mathematical learning.

References

148AppsBiz. (2015). *App store metrics*. Retrieved from http://148apps.biz/app-store-metrics/

Attard, C., & Curry, C. (2012). Exploring the use of iPads to engage young students with mathematics. In J. Dindyl, L. P. Cheng, & S. F. Ng (Eds.), *Mathematics education: Expanding horizons* (Proceedings of the 35th annual conference of the Mathematics Education Research Group of Australasia, Singapore, pp. 75-82). Singapore: MERGA.

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

Larkin, K. (2015 in press.). "An App! An App! - My Kingdom for an App": An 18 month quest to determine whether apps support mathematical knowledge building. In T. Lowrie & R. Jorgensen (Eds.), *Digital games and mathematics learning: Potential, promises and pitfalls.* Springer Press.

Bos, B. (2009). Virtual math objects with pedagogical, mathematical, and cognitive fidelity. *Computers in Human Behavior, 25*, 521-528. doi: 10.1016/j.chb.2008.11.002

Burns, B. A., & Hamm, E. M. (2011). A comparison of concrete and virtual manipulative use in third and fourth grade mathematics. *School Science and Mathematics, 111*(6), 256-261. doi: 10.1111/j.1949-8594.2011.00086.x

Carbonneau, K. J., Marley, S. C., & Selig, J. (2013). A meta-analysis of the efficacy of teaching mathematics with concrete manipulatives. *Journal of Educational Psychology, 105*(2), 380-400. doi: 10.1037/a0031084

Dick, T. P. (2008). Fidelity in technological tools for mathematics education. In G. Blume & M. Reid (Eds.), *Research on Technology and the Teaching and Learning of Mathematics: Cases and Perspectives* (Vol 2, pp. 333-339). Information Age Publishing.

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

Moyer-Packenham, P. S. et al. (2015). Young children's learning performance and efficiency when using virtual manipulative mathematics iPad apps. *Journal of Computers in Mathematics and Science Teaching, 34*(1), 41-69.

Özel, S., Özel, Z., & Cifuentes, L. (2014). Effectiveness of an online manipulative tool and students' technology acceptances. *International Journal of Educational Studies in Mathematics, 1* (1), 1-15.

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. K. Lester (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 1169-1207). Charlotte, NC: Information Age.