

Cite as: Lee, C-Y. & Cherner, T. S. (2015). A comprehensive evaluation rubric for assessing instructional apps. *Journal of Information Technology Education: Research*, 14, 21-53. Retrieved from <http://www.jite.org/documents/Vol14/JITEV14ResearchP021-053Yuan0700.pdf>

A Comprehensive Evaluation Rubric for Assessing Instructional Apps

Cheng-Yuan Lee and Todd Sloan Cherner
Coastal Carolina University, Conway, SC, USA

clee@coastal.edu tcherner@coastal.edu

Abstract

There is a pressing need for an evaluation rubric that examines all aspects of educational apps designed for instructional purposes. In past decades, many rubrics have been developed for evaluating educational computer-based programs; however, rubrics designed for evaluating the instructional implications of educational apps are scarce. When an Internet search for existing rubrics was conducted, only two such rubrics were found, and the evaluation criteria used in those rubrics was not clearly linked to previously conducted research nor were their evaluative dimensions clearly defined. These shortcomings result in reviewers being unable to use those rubrics to provide teachers with a precise analysis of an educational app's instructional potential. In response, this paper presents a comprehensive rubric with 24-evaluative dimensions tailored specifically to analyze the educational potential of instructional apps.

Keywords: evaluation rubric, tablet technology, instructional apps, tablet devices, apps, public education, blended learning

Introduction

With the current explosion of tablet computing devices, school districts from across the nation are purchasing tablets for all their students and teachers to use (Murray & Olcese, 2011; Pilgrim, Bledsoe, & Reily, 2012). To ensure teachers are using these devices effectively, school districts have adopted Blended Learning instructional models (Graham, 2005; Picciano & Seaman, 2007). The challenge, however, is that with over 20,000 iOS educational apps available in the App Store (Earl, 2013; Rao, 2012), teachers need support in identifying quality apps to use or they risk wasting their time with inferior apps. This situation creates a pressing need for an evaluation rubric that examines the quality of educational apps. With such a rubric, teachers would have a tool that supports them in identifying quality apps to use with their students as part of their blended learning instruction, and educational app designers would have indicators to consider when developing their apps.

Material published as part of this publication, either on-line or in print, is copyrighted by the Informing Science Institute. Permission to make digital or paper copy of part or all of these works for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage AND that copies 1) bear this notice in full and 2) give the full citation on the first page. It is permissible to abstract these works so long as credit is given. To copy in all other cases or to republish or to post on a server or to redistribute to lists requires specific permission and payment of a fee. Contact Publisher@InformingScience.org to request redistribution permission.

Over the past few decades, many rubrics have been developed for evaluating educational computer-based programs (Coughlan & Morar, 2008; Elissavet & Economides, 2003; Kennedy, Petrovic & Keppell, 1998; Schibeci, Lake, Philips, Lowe, Cummings, & Miller, 2008; Shiratuddin & Landoni, 2002); yet, rubrics that focus specifically on educational tablet apps are extremely

Editor: Lorraine Staehr

Submitted 23 May 2015, Revised 16 December 2015, Accepted 1 January 2015

limited. When an Internet search for existing rubrics was conducted, only two such rubrics were identified (Buckler, 2012; Walker, 2010). The problem is that the evaluation criteria used in these rubrics is not clearly linked to previously conducted research nor are the rubrics' evaluative dimensions clearly defined. These shortcomings do not allow reviewers to provide teachers with precise analyses of an instructional app's educational potential. Instructional apps provide students with a learning experience that results in students acquiring a skill or learning academic information, and these apps can be classified as skill-based, content-based, or function-based (Cherner, Dix, & Lee; 2014). In response, this paper presents a new comprehensive rubric with 24-evaluative dimensions designed specifically to analyze instructional apps designed for tablet devices.

Background of Educational Software Evaluation Rubrics

Because computers have been used increasingly more for educational purposes since the 1980s (Kulik & Kulik, 1991; Papert, 1993), teachers have often been faced with the challenge of selecting appropriate computer software for their students to use. As a teacher, knowing how to evaluate a computer software product is as important as knowing how to use it (Winslow, Dickerson, & Lee, 2013). Due to a huge repository of software applications being available that includes CD-ROMs, the Internet, and mobile devices, educators and researchers have developed multiple frameworks for evaluating educational computer software to support teachers' selection of quality software, with Reeves and Harmon's (1993) *Systematic Evaluation of Computer-Based Education* being widely recognized and adapted (Cronjé, 2006; Ehlers & Pawlowski, 2006; Elissavet & Economides, 2003; Phillips, 2005; Schibeci et al., 2008). In their framework, Reeves and Harmon put forward 14 pedagogical evaluative dimensions and 10 user-interface evaluative dimensions, and these dimensions have served as a foundation for many rubrics developed in the last two decades (Coughlan & Morar, 2008; Elissavet & Economides, 2003; Kennedy et al., 1998; Schibeci, et al., 2008; Shiratuddin & Landoni, 2002).

As the rise of mobile technology (e.g., smartphones and tablet technologies) in the late 2000s has swept the globe, it represents a paradigm shift in how technology is being used in schools (Leung & Chan, 2003). These mobile devices are so highly portable and powerful that students and teachers can use them anytime and anywhere. As a result, school districts from across the nation are purchasing tablet technologies for all their students and teachers to use (Murray & Olcese, 2011; Pilgrim et al., 2012), and software companies are developing applications, commonly referred to as "apps," to be loaded onto these mobile devices for educational purposes. Currently, with a huge amount of educational apps available for download (Earl, 2013; Rao, 2012) and more apps being continually developed, teachers need an evaluation framework to support them in selecting quality apps. However, Reeves and Harmon's (1993) original evaluation model is insufficient for assessing educational apps because:

1. Webb's (1997) *Depth of Knowledge* (DoK) and *21st Century Skills* go unaddressed in their model. With the Common Core State Standards, educators need to deeply consider the rigor of the learning tasks that software requires of students and how completing those tasks better prepares students to compete in the global marketplace (Hess, Carlock, Jones, & Walkup; 2009).
2. Some of Reeves and Harmon's evaluative dimensions are too vague or too broad, which makes them not measurable. Examples of these dimensions include origin of motivation, mapping, and overall functionality. As technology continues to evolve, new and updated language is needed to assess these finer points, especially in regards to instructional apps.
3. In Reeves and Harmon's model, indicators that provide accurate ratings for each evaluative dimension were not provided. Rather, Reeves and Harmon only offered a brief description about an aspect of software that could be evaluated. They did not provide a ru-

bric complete with dimension headings and indicators to support how reviewers can evaluate educational software.

4. Reeves and Harmon's model does not address some of the new functionalities that tablet technology utilizes. Examples of these functionalities include sharing capability, cross-platform integration, and ability to save progress.

While Reeves and Harmon's (1993) evaluative dimensions provide a strong foundation, revising and evolving their original work is needed for this post-PC age (Greenfield, 2006). To ensure an instrument that addresses these concerns has not already been created, a review of literature was conducted.

To conduct this review, a comprehensive search was performed with *Google Scholar*, *ERIC*, and *EBSCOhost*, using "rubric", "education software", "multimedia", "usability test", "interactivity", and "interface" as search keywords to locate research studies relevant to the evaluation rubric for assessing educational apps. Although several hundred articles were yielded from the search results, the majority of the articles addressed the evaluation of computer-based applications, websites, or CD-ROM, and only two articles (e.g. Buckler, 2012; Walker, 2010) employed rubrics to assess the quality of educational apps. However, the rubrics they put forward are not comprehensive and include specific shortcomings.

Buckler (2012) created a rubric with six dimensions to evaluate apps for use with adults who have special needs, and the domains he used included application, feedback, adjustability, ease of use, cost, and benefits. Walker (2010) developed another rubric with six dimensions for evaluating the quality of apps, and his dimensions included curriculum connections, authenticity, feedback, differentiation, user friendliness, and motivation. Upon careful review of these two rubrics, multiple important factors were not considered or were not fully articulated, which include:

1. *The rubrics were not directly linked to research.* During the 1990s, a significant amount of knowledge for evaluating the educational potential of CD-ROMs was developed (Herrington & Oliver, 1997; Kennedy et al., 1998; Overbaugh, 1994; Reeves, 1994; Reeves & Harmon, 1993; Reeves & Hedberg, 2003). When reviewing Buckler's and Walker's rubrics, explanations for how their rubrics were grounded in or informed by previously conducted research and best practices were omitted. Therefore, drawing conclusions about their rubrics' validity is challenging.
2. *The rubrics were not fully developed.* Buckler's and Walker's rubrics only touched on surface-level evaluations of apps. For example, Walker's rubric makes general statements about why an app may be motivating to use, but there are multiple considerations that go unaddressed. These considerations include how easy an app is to use, the pace of material presented to the learner, the content of the app, and the learner's motivation for using the app among several other possible factors (Pintrich, 2003; Reeves, 1994; Reeves & Harmon, 1993; Reeves & Hedberg, 2003). By lumping these considerations under the general "Student Motivation" dimension, Walker's rubric does not take into consideration the finer points of why a student may or may not be motivated to use an app.
3. *The specific terms used in these rubrics are limited.* For example, in Walker's description of *differentiation*, he explained it as "the ability to set the level of difficulty or target specific skills for individual children [that] increases the usefulness of the app as an instructional tool" (2010, p. 5). The problem is that by limiting differentiation to selecting the difficulty level of an app's content or specific skills taught, Walker is ignoring multiple components of differentiated instruction such as the learner's background knowledge, learning style, and interest (Tomlinson, 1999, 2001). By limiting his use of *differentiation* to the content or skill taught by an app, he is supporting a reductionist view of differentiated instruction and how differentiated instruction benefits different types of learners.

4. *The two rubrics each use a 4-point system to evaluate the quality of apps.* Whereas traditional Likert Scale models use a 5-point scale (Jamieson, 2004; Matell & Jacoby, 1971; Trochim, 2006), the 4-point scale used by both Buckler and Walker is limiting in that it does not allow for more nuanced distinctions to be made about the apps it is assessing. The more detailed data a rubric provides teachers about an app, the more informed and effective teachers will be when using the app in their classroom (Winslow et al., 2013).
5. *The rubrics are not comprehensive.* Because each of the rubrics only uses six dimensions to evaluate apps, they do not allow for thorough evaluations of apps. For example, detailed analyses of an app's design, instructional value, and potential to engage learners are not individually addressed. Rather, these concerns are grouped together in the general dimensions used by both Buckler and Walker. These generalized groupings, therefore, limit the type of analyses that can be done and the implications about an app's strengths and weaknesses derived from those analyses.

If teachers are expected to integrate apps into their teaching practice, they need a valid and research-based tool to analyze the quality of apps they wish to use in their classroom. For the reasons listed above, major concerns about the rubrics put forward by Buckler and Walker exist, which creates the need for a new rubric to evaluate instructional apps designed for educational purposes.

The Creation of a New App Evaluation Rubric

Based on previously conducted and published works about methods used to evaluate technologies for educational purposes (Buckler, 2012; Coughlan & Morar, 2008; Pintrich, 2003; Premkumar, Hunter, Davison, & Jennett, 1998; Reeves, 1994; Reeves & Harmon, 1993; Squires & Preece, 1999; Walker, 2010), a comprehensive rubric with 24-evaluative dimensions was developed for assessing the quality of instructional apps (see Appendix A). Table 1 synthesizes the relevant literature that laid the solid foundation for the development of the new 24-dimension rubric.

After these 24 dimensions were developed, they were aligned to a 5-point Likert scale format so quantitative measures could be established for each dimension. In addition, indicators for the 1-5 point ratings for each dimension were clearly stated to avoid as much ambiguity as possible in the rating process.

To verify its face and content validities (Haynes, Richard, & Kubany, 1995; Lynn, 1986), the initial rubric was examined and critiqued by two groups of experts. The first group consisted of four experienced faculty members who were actively involved in teaching at the university level and held a doctoral degree in the field of Educational Technology, Educational Psychology, or Literacy. The second group included public school teachers and graduate students who held professional teaching licenses and had at least five years classroom teaching experience. The experts from each group were asked to comment on each dimension of the rubric and rate the dimension with the assessment form included in Appendix B. Modifications were then made to the rubric based upon the experts' suggestions. After the revised rubric was finalized, it was presented to the experts again to ensure that they understood each of the rubric's dimensions and indicators. After reviewing the revised rubric, all the experts confirmed that its clarity was increased and that they understood each of the rubric's dimensions and indicators.

Table 1: Theoretical Framework Related to Rubric Dimensions

Theoretical Framework	Relevant Literature	Rubric Dimensions
Common Core State Standards for Literacy	Fisher & Frey, 2013	A3. Connections to Future Learning C7. Utility
Cooperative learning	Blumenfeld, Marx, Soloway, & Krajcik, 1996; Gokhale, 1995; Slavin, 1990a, 1990b, 1992	A7. Cooperative Learning
Human Interactive design	Duffy & Jonassen, 1992; Elawar & Corno, 1985; Hannafin, 1992; Huba & Freed, 2000; Kearsley, 1988; Ovando, 1994; Palloff & Pratt, 1999; Reeves & Harmon, 1993; Reeves & Hedberg, 2003; Reeves & Reeves, 1997	A4. Value of Errors A5. Feedback to Teacher A8. Accommodation of Individual Differences B5. Navigation B7. Information Presentation B8. Media Integration B9. Cultural Sensitivity C2. Interactivity C6. Aesthetics
Human motivation	Eccles & Wigfield, 1995; Pintrich & Schunk, 1996	C5. Interest C6. Aesthetics C7. Utility
Multimedia learning principles	Mayer & Anderson, 1991, 1992; Mayer & Moreno, 1998	B8. Media Integration
Rigor and relevancy taxonomies	Anderson, Krathwohl, & Bloom, 2001; Dagget, 2005; Webb, 1997; 1999	A1. Rigor A2. 21 st Century Skills
Usability design	Coughlan & Morar, 2008; Elissavet & Economides, 2003; Galitz, 1985, 1992; Kearsley, 1988; Kennedy et al., 1998; Lucas, 1991; Schibeci, et al., 2008; Sherr, 1979; Shiratuddin & Landoni, 2002	B3. Screen Design B4. Ease of Use C2. Interactivity
Zone of proximal development	Vygotsky, 1978	A6. Level of Learning Material

To assist teachers in using this comprehensive rubric, the 24 dimensions were categorized into three domains: (A) *Instruction*, (B) *Design*, and (C) *Engagement*. In the following sections, the evaluative dimensions contained in each domain are introduced and discussed in detail. However, a few key considerations about using this rubric need to be clarified. First, each of the 24 evaluative dimensions together comprise the entire rubric that is located in Appendix A, and the dimensions being discussed in the following sections refer back to it. Second, each evaluative dimension was designed to follow a consistent format. The format includes a *prompt* that focuses the dimension on a central question, and five *indicator descriptors* that describe the ways in which an app's functionality or design may behave in response to a prompt. Moreover, the descriptors are ranked using a 5-point Likert Scale, with 5 being the highest quality and 1 being the lowest quality.

ty. Third, an instructional app that is representative of the type of app that would score highly in each dimension is discussed. This discussion is provided to exemplify the characteristics of an app that could earn a high score. Fourth, a *not applicable* (N/A) option exists as part of each dimension. This option was included because some instructional apps by design are not assessable on all dimensions. For example, an app that is a collection of poems and is designed for learners to read poetry may not include an assessment component that evaluates comprehension of the poems read. Therefore, because the app is a content-based app (Cherner et al., 2014), it was purposely designed to allow learners access to the knowledge, texts, or information about a specific topic. It was not designed to evaluate the knowledge or information learners gained from reading the app's different texts. As such, dimensions that were designed to assess learner comprehension are rated as N/A because they were not part of the app's purpose. Finally, each of this rubric's dimensions was designed specifically to measure only instructional apps. Educational apps designed to support teachers in other ways such as planning instruction, assessing student work, and tracking attendance fall outside the scope of this rubric. Therefore, the word "app" in this rubric only refers to instructional apps.

Domain A: Instruction

The instructional worth of an app is the rubric's first domain, and its dimensions were designed to measure an app's educational value. An app's educational value worth is defined as an analysis of the cognitive demands placed on learners and the support offered to them by an app as they work to meet its learning objective. An app's learning objective is the knowledge or skills learners are supposed to acquire or strengthen by engaging it. Eight dimensions are used to evaluate an app's instructional worth. These include (A1) Rigor, (A2) 21st Century Skills, (A3) Connections to Future Learning, (A4) Value of Errors, (A5) Feedback to Teacher, (A6) Level of Material, (A7) Cooperative Learning, and (A8) Accommodation of Individual Differences.

A1. Rigor

The "Rigor" dimension measures the thinking skills an app requires of learners as they engage it. To frame this dimension, Webb's (1997) DoK was selected because of previous researchers' use of it to analyze the rigor of learning tasks and different types of assessments (Blackburn, 2014; Herman & Linn, 2013; Webb, 1999). The DoK is divided into four levels, and each represents a specific type of rigor. At Level 1, an app requires learners to recall only information to complete a problem or answer a question. At Level 2, an app requires learners to use knowledge and skills when classifying, comparing, and summarizing different concepts and terms to complete a learning task. At Level 3, an app requires learners to think strategically when assessing, formulating, and making logical statements about ideas and phenomena. At Level 4, which is the most rigorous, an app requires learners to extend their thinking by creating, critiquing, and synthesizing information to complete a learning task. With this measure, evaluators are assessing the thinking skills learners engage when completing a learning task required by an app.

Example App for Rigor. Apps that allow learners to create a learning artifact to explain, assess, or critique a topic score well on this dimension. For example, *Touchcast* lets learners create high-quality videos that can be used to explain, evaluate, or offer multiple ideas about any concept, law, or phenomenon. The videos learners can create using *Touchcast* extend how they understand, engage, and explain the topic of their video. Because learners are creating these videos and the implications for how these videos can be used, this app's rigor is rated as a 5 according to this dimension's indicators.

A2. 21st Century skills

The “21st Century Skills” dimension analyzes the types of skills learners use while engaging an app to see if it prepares them for the 21st Century, technology-enhanced, modern world. To define specifically the abilities that are considered “21st Century” skills, materials released by multiple professional organizations including The Institute of Museum and Library Services (2014), Partnership for 21st Century Skills (2009), and Assessment & Teaching of 21st Century Skills (2014) were consulted. After reviewing these materials, the different abilities identified were organized into four main categories that include the ability to (1) solve complex problems, (2) collaborate and communicate with peers, (3) use technology effectively, and (4) be an informed, global citizen. With these guiding principles, apps can be evaluated regarding how they prepare learners for success using 21st Century Skills.

Example App for 21st Century Skills. Apps that allow learners to collaborate to complete learning tasks score well on this dimension. For example, *Google Drive* and apps that function as digital classrooms (e.g., *Edmodo* and *Schoology*) are rated highly because they offer learners the opportunity to use technology together as a tool for solving complex problems about a range of issues. As learners use technology, they must communicate to plan how they will complete the learning activity. Because teachers can create learning activities that require higher-order thinking skills to solve challenging, relevant problems, these apps require students to use 21st Century Skills, and they are each rated as a 5 on this dimension.

A3. Connections to future learning

The “Connections to Future Learning” dimension explores if an app prepares learners for the literacy and numeracy-oriented tasks that will be placed on them. As represented by the Common Core State Standards (National Governors Association, 2010), the C3 Framework (National Council for Social Studies, 2013), and the Next Generation Science Standards (Next Generation Science Standards Lead States, 2013), the literacy and numeracy skills learners must possess to be “College- and Career-Ready” (National Center on Education and the Economy, 2013) are changing. No longer can skills be taught in isolation of other skills; rather, the skills learners acquire must build on each other so learners are prepared to enter a post-secondary educational program or the workforce. In response, this dimension gives consideration to how an app teaches or reinforces specific literacy and numeracy skills for future learning.

Example App for Connections to Future Learning. Apps that teach foundational knowledge and skills score higher on this dimension than apps that review previously learned knowledge and skills. For example, *Cargo-Bot* exemplifies the characteristics of an app that prepares learners for future learning. Using a game format to teach learners basic computer programming skills, *Cargo-Bot* builds learners’ foundational computer literacy skills, which they will need if they enter the field of computer science. Because *Cargo-Bot* teaches learners the foundational skills needed for a specific literacy, it scores a 5 on this dimension.

A4. Value of errors

The “Value of Errors” dimension investigates how an app allows learners to make mistakes and learn from their experience. Constructive feedback and individualized instructions have long been recognized as key characteristics for effective teaching (Elawar & Corno, 1985; Ovando, 1994; Palloff & Pratt, 1999). When computing technology was introduced to multimedia production during the 1990s, immediate feedback and adaptation to learners’ level of mastery were made possible. In relation to tablet technologies, apps that offer high-quality feedback provide useful information to learners about their progress, which supports them as they work to gain new skills and knowledge.

Example App for Value of Errors. Apps that provide specific feedback about why a learner's answer was incorrect and offer the learner an opportunity to answer the question again based on that feedback score well on this dimension. For example, when learners use the *QuotEdReading* app, it teaches them general literacy skills by first presenting a reading passage to them before asking them follow-up questions about the passage's main idea(s), details, and vocabulary. If learners answer correctly, they progress to the next question. If learners answer incorrectly, *QuotEdReading* explains why they answered incorrectly before giving them an opportunity to answer it again. In this way, learners are able to use the feedback provided by this app, and it scores a 5 on this dimension.

A5. Feedback to teacher

The "feedback to teacher" dimension evaluates if and how an app allows teachers to monitor their learners' progress. In any form of technology-based learning, the role of teacher has gradually shifted from being didactic to that of facilitator (Berge, 1995; Collins, 1991; Jeffries, 2005; Reeves, 1994; Reeves & Reeves, 1997). In order for teachers to assume the role of facilitator, they need to know the content their learners studied and the mistakes they made, which allows them to track their learners' progress. That information then allows teachers to craft future instruction to meet the specific needs of their learners. To be considered effective, an app must supply teachers with data about their students' performance.

Example App for Feedback to Teacher. Apps that allow teachers to easily monitor the progress their learners make while engaging its lessons, tutorials, and/or instructional activities score highly on this app. For example, *Blackboard Madness: Math* is designed to develop learners' numeracy skills, and it provides teachers with reports about the amount of instructional activities their learners complete. Specific information recorded in this app includes the total amount of time learners spent playing the games, the different games they played, how many rounds they played, their average scores per round, and any specific achievements they earned. These data are all saved and available for teachers to analyze, so they can design lessons to meet their learners' specific needs based on it. However, because that data is only available in the app and cannot be sent over email or accessed on a website, *Blackboard Madness: Math* scores a 3 on this dimension.

A6. Level of learning material

The "Level of Learning Material" dimension evaluates if an app's material is appropriate for its target group of learners, and "learning material" is defined as the content or activities learners engage to acquire a skill or gain understanding of a topic. For any learning material, clearly stating its target audience is critical to its potential learners. Mismatches between learners' zone of proximal development (Vygotsky, 1978) and the difficulty level of the learning material they engage will likely lead to unfavorable learning results. Easy tasks presented to advanced or experienced learners leads to boredom and loss of interest; whereas, difficult learning materials given to beginning learners result in frustration and anxiety (Csikszentmihalyi, 1997; Juul, 2004; Rieber, 1996). Correctly matching learners' ability level to the learning materials being provided must be a consideration when an app is being developed. This requires that an app is designed for a target group of learners and its learning materials are aligned to that group. To score highly on this dimension, the app must ensure that its content is cognitively and developmentally appropriate for its target group of learners.

Example App for Level of Learning Material. Apps that provide different levels of content based on the complexity of the knowledge or information being taught score well on this dimension because learners are given the option to choose the content that is most appropriate for them. For example, *One Minute Reader*, which is designed to develop learner's reading fluency and comprehension skills, gives learners multiple levels of content to select. Plus, it uses different built-in

reading assessments to help learners choose the content that is appropriate for them, and the content is listed by grade level. In this way, *One Minute Reader* scores a 5 on this dimension because the content learners select is organized by grade-level and is appropriate for students in that grade level.

A7. Cooperative learning

The “Cooperative Learning” dimension evaluates if an app provides learners the opportunity to work on group projects together. Multiple studies have proven that cooperative learning benefit learners instructionally and socially (Blumenfeld et al., 1996; Gokhale, 1995; Slavin, 1992). When learners are given an opportunity to work together in pairs or small groups to accomplish shared goals, they tend to perform better and their social skills also improve (Johnson & Johnson, 1987; Slavin, 1990a, 1990b). Currently, most mobile devices are equipped with Wi-Fi, Bluetooth, and the Internet. As such, this dimension analyzes if an app’s instructional activities are designed in a way that lets learners collaborate on projects and achieve shared learning goals.

Example App for Cooperative Learning. Apps that let multiple learners view, share, and collaborate on projects together score highly on this dimension, and several apps allow for learners to view and share projects using a cloud storage system, sending them over email, or uploading them to a website. However, apps can also be designed for learners to create projects together, and *Educreations* is designed for that type of collaboration. *Educreations* lets learners create and share presentations, and learners can work together when planning, developing, and recording their presentations. Once a lesson is created, learners can post it to *Educreations*’ online forum, where it can be viewed by others. In this way, learners are collaborating in order to create their presentation and sharing it online where it can be accessed by other learners. However, learners are limited in how they can communicate with other learners because they can communicate only within the contents of their presentation. For these reasons, *Educreations* scores a 4 on this dimension.

A8. Accommodation of individual differences

The “Accommodation of Individual Differences” dimension investigates whether an app is designed to accommodate a wide range of learners. Students enter the classroom with differences in their aptitudes, background knowledge, language proficiency, motivation, cultural heritage, and learning styles among other diversities (Jonassen & Grabowski, 2012; Reeves, 1994; Tomlinson, 1999). When measuring the effectiveness of any computer-based instruction, accommodation of individual differences is a critical factor (Reeves & Harmon, 1993). An app with a high rating in this dimension customizes its content for different learners based on their specific needs, which supports the needs of diverse learners.

Example App for Accommodation of Individual Differences. Apps that allow learners to choose the content they engage, offer different methods for engaging that content, and build background knowledge about the content score highly on this dimension, and *Color Uncovered* is an app that includes these characteristics. With *Color Uncovered*, learners are able to explore a range of phenomena related to color by selecting from a variety of topics about color that may interest them. Next, the app allows learners to read about the color phenomenon to build their background knowledge before letting them complete an activity about the phenomenon, and learners have the option to skip this reading. Most of *Color Uncovered*’s activities have learners view the phenomenon and then respond to it by tapping their screen to see a new component of the phenomenon or by doing a physical movement to experience the phenomenon firsthand. Because learners can choose the different topics to engage and the variety of ways learners may engage those topics, *Color Uncovered* scores a 4 on this dimension.

Domain B: Design

After examining an app's instructional worth, the next consideration is its functionality. A high-quality app not only provides substantial instructional value to learners, but it also allows them to interact smoothly and efficiently with it. Thus, the second domain assesses the functionality of an app's design, and functionality used in this context is defined as an analysis of how learners interact with an app. This domain is comprised of nine dimensions that measures an app's design according to its (B1) Ability to Save Users' Progress, (B2) Platform Integration, (B3) Screen Design, (B4) Ease of Use, (B5) Navigation, (B6) Goal Orientation, (B7) Information Presentation, (B8) Media Integration, and (B9) Cultural Sensitivity.

B1. Ability to save progress

The "Ability to Save Progress" dimension explores if an app allows learners to save their progress and then resume their activities at a later time. This is the first new dimension that was not introduced in any of the existing rubrics. When a learning process is interrupted and then is resumed at a later time, learners should not have to start from the beginning. Instead, they should be able to resume engaging the content from where they previously left off, and the app should allow them to do so automatically. For example, if a child is learning the alphabet and only gets through the letter "G" in one sitting, that child should not have to begin with the letter "A" when returning to the app. Rather, that child should be able to start with the letter "H." Allowing learners to return to the content they were last engaging before logging off lets them use their time more efficiently, and it ultimately increases their learning opportunities.

Example App for Ability to Save Progress. Apps that quickly and automatically load the content learners were last engaging before they logged off score well on this dimension. For example, *Art HD* is an app that consists of a gallery of renowned artists' works. This app's organizational design allows learners to browse the artists' works alphabetically or by movement and period. When learners finish browsing the works and close the app, they can reopen it and be returned automatically to the work they were last viewing, which results in *Art HD* being rated as a 5 on this dimension.

B2. Platform integration

The "Platform Integration" dimension analyzes if an app connects its content to different platforms, and this is the second new dimension that goes unaddressed in any existing rubric. The term *Platform Integration* refers to any artifacts or documents stored either in an app or created by using an app that can be shared or opened in other platforms. These platforms include other apps, online communities, independent websites, and learners' email. This dimension is important because being able to share learning artifacts and content across platforms adds accessibility to the learning artifacts created and knowledge gained. Because sharing content across platforms is a relatively new design feature, it went unaddressed by previously created rubrics. However, due to its implications for educational uses, it is included in this rubric.

Example App for Platform Integration. Apps that allow their content or learner created work to be shared across multiple platforms score highly on this dimension, and *Bamboo Paper* includes these features. With *Bamboo Paper*, learners can create documents that can be shared via email, viewed in other apps, and stored online. Because the created documents can be shared in these three different capacities, but not in an online community, it scores a 4 on this dimension.

B3. Screen design

The "Screen Design" dimension analyzes if an app's text, graphics, videos, sound, and speech are well organized. Multiple researchers (Galitz, 1985, 1992; Lucas, 1991; Sherr, 1979) have stated

that screen design is a critical factor for human-machine interaction. An effective interface between a learner and a computer facilitates effective communication, which leads to a better learning performance. Reeves and Harmon (1993) proclaimed that screen design is a complex dimension that can be broken down into many elements such as text, icons, graphics, color, and other visual aspects of computer programs. However, it is too trivial to measure each of those elements separately. As a result, in this rubric, screen design is treated holistically by considering if an app's text, graphics, videos, sound, and speech are organized in a way that enhances an app's purpose.

Example App for Screen Design. Apps that integrate the media to enrich how learners experience its content score highly on this dimension, and *Louvre HD* exemplifies these characters. *Louvre HD* houses a collection of art that learners can search by artist's name, genre, or time period. As learners peruse the works of art, they can listen to classical music while reading the artists' name and title of different works, and the way this app combines its visual and audio elements always enhances its content. For this reason, *Louvre HD* is rated as a 5 on this dimension.

B4. Ease of use

The "Ease of Use" dimension assesses if an app is intuitive and user-friendly, and this dimension has appeared repeatedly in multiple evaluation rubrics (Coughlan & Morar, 2008; Elissavet & Economides, 2003; Kennedy et al., 1998; Schibeci et al., 2008; Shiratuddin & Landoni, 2002). This dimension is important because learners who find an app easy to use are more likely to engage it. If an app is difficult to use, learners are not likely to spend time with it. Plus, with so many other alternative apps available, putting forth energy to understand how to use an app can be a frustrating experience. Therefore, it is essential that apps are intuitive and can be easily engaged by learners the first time they are being used.

Example App for Ease of Use. Apps that are self-explanatory and learners can immediately engage without guidance score best on this dimension. For example, *Peterson Backyard Birds – A Field Guide to Birds of North America* is an app that learners can use quickly and easily. This app is a reference resource learners can use to identify different birds that live in North America. This app's interface design is intuitive, and it allows learners to quickly utilize and engage all of its content, functions, and features without needing additional assistance or guidance. For these reasons, this app scores a 5 on this dimension.

B5. Navigation

The "Navigation" dimension looks at how learners move through an app's content and options. A major aspect of navigation is orientation, meaning that learners know where they are and where they can go in a computer program (Reeves & Harmon, 1993). An app with a clear navigational system in place supports learners as they move through an app's content, which helps them maintain their orientation in the program. These supports may include an app having breadcrumbs, using navigational arrows, or providing a map of the app's content. With these supports in place, learners are more likely to put energy towards engaging the app's content and learning goals rather than losing their place in a program and becoming frustrated.

Example App for Navigation. Apps that provide clear, easy-to-use tools that learners can utilize when moving through content score highly on this dimension. For example, the *NASA app*, which contains images, recordings, and text about the United States' space exploration program, includes tools for learners to use when moving through its content. These tools include navigational buttons that are big and located consistently on each screen. Without much effort and in a short time period, learners will know how much content is available to them in the app and be able to explore it using these buttons. Because its navigational buttons are intuitive to use and presented

in a consistent format, it results in learners being able to easily navigate the *NASA App*. In this way, it scores a 5 when rated using this dimension.

B6. Goal orientation

The “Goal Orientation” dimension analyzes whether each component of an app contributes to learners’ meeting the instructional objective set for them. According to Coughlan and Morar (2008), an effective piece of educational software ensures that its content and activities are aligned to its instructional objective(s), and “instructional objective” in this context is defined as the skills learners are to develop or knowledge learners are to acquire by using an app. Each time it deviates from its instructional objective, the app’s educational worth drops. Therefore, every app designed for educational purposes should have its instructional objective aligned tightly to its content.

Example App for Goal Orientation. Apps that match their purpose with their content score highly on this dimension, and *Star Chart* provides an example of an app that has its content tightly aligned to its instructional objective. The instructional objective of *Star Chart* is for learners to gain knowledge about the different constellations. To support learners in gaining this knowledge, different activities, graphics, sounds, and animations are included in this app, and each of those resources facilitates learners meeting this app’s instructional objective. Additionally, there are no distractions included in this app that may impede learners from realizing its instructional objective. As such, *Star Chart* is rated a 5 on this dimension.

B7. Information presentation

The “Information Presentation” dimension is concerned with learners’ ability to understand and gain knowledge by the way an app presents information to them. Imagine a well-written paper with no proper formatting: no title, no headings, and no paragraph breaks. The paper is rich in content, but its lack of text features makes learning from it challenging, and the same is true for apps. If an app presents information in a way that is incomprehensible to learners, the app’s entire instructional value is significantly reduced, no matter how rich and meaningful the information is or how elegant its interface is (Reeves & Harmon, 1993). Apps must use a clear presentation format, so learners can deeply engage its content. Apps that rate highly in this dimension present their content using a logical format learners can follow easily, which supports learning in achieving an app’s learning goals.

Example App for Information Presentation. Apps that have a well-organized main screen and use a consistent format for its other screens score highly on this dimension, and *The Poetry App* exemplifies these qualities. To explain, *The Poetry App* contains multiple ways for learners to engage poetry. With this app, learners can listen to and read famous poetic works, read poets’ biographies, and create their own poetry. Due to multiple avenues for how learners can engage its content, this app organized its home screen in a way that clearly indicates the type of content learners will engage when they select a specific option. Plus, the other screens are formatted in a consistent way for learners, which results in this app being rated as a 5 on this dimension.

B8. Media integration

The “Media Integration” dimension analyzes how well an app combines its different media components (e.g., text, graphics, videos, sound, and speech) to form a cohesive program (Reeves & Harmon, 1993). The various media components in an app should work together to form a quality instructional experience for learners (Mayer & Anderson, 1991, 1992; Mayer & Moreno, 1998). If these components are not integrated with each other, it can result in a disjointed program. A high score in this dimension indicates that the media components support each other, which results in the app providing learners with a cohesive program to use.

Example App for Media Integration. Apps that combine different media components to form a smooth experience for learners score highly on this dimension. For example, the *Music Tutor Free (Sight Reading Improver)* app effectively weaves various note sounds, music, and graphics together into one elegant program, which supports learners acquiring the skills they need to read music. Because of the way the audio and visual features complement one another and support learners meeting this app's instructional objective, it scores a 5 on this dimension.

B9. Cultural sensitivity

The "Cultural Sensitivity" dimension analyzes an app's ability to accommodate learners' diverse ethnic and cultural backgrounds (Reeves & Harmon, 1993). Although every piece of instruction may not be able to adapt to every cultural norm, educational software should be designed to be as culturally sensitive as possible (Powell, 1993). In regards to apps, cultural sensitivity requires them to promote the spreading of authentic, accurate information about a culture and not stereotypical or biased depictions. When cultural sensitivity is taken into consideration, learners of all backgrounds can access educational apps easily and safely, which mitigates fear of encountering insensitive content or hurtful representations.

Example App for Cultural Sensitivity. Apps that use media for learners to view, hear, or communicate with diverse cultures score well on this dimension. For example, the *YouTube* app allows learners to search and identify videos created by people from across the globe that accurately portray their culture. By viewing videos uploaded by individuals from around the globe, *YouTube* offers learners an opportunity to connect with diverse cultures in an authentic manner. Because *YouTube* allows learners to connect with multiple cultures and the videos can contain authentic images of different cultures, it scores a 5 on this dimension.

Domain C: Engagement

If an app is ranked high for its educational value and functionality but fails to motivate its learners to interact and learn from it, it does not possess much value to the learners. Because specific reasons why one learner may engage an app more than another learner are beyond the scope of this rubric in general and this domain in particular, the dimensions included here offer representative reasons for why an app may appeal to learners. The Engagement domain is comprised of seven dimensions, and each dimension was created to measure how motivated learners might be to interact with and learn from an app. The dimensions used to evaluate if apps are likely to appeal to learners include (C1) Learner Control, (C2) Interactivity, (C3) Pace, (C4) Personal Preferences, (C5) Interest, (C6) Aesthetics, and (C7) Utility.

C1. Learner control

The "Learner Control" dimension analyzes if an app allows learners to select the level they engage its content or paths they follow through interactive material (Reeves, 1994). It is often assumed that more learner control results in higher motivation of learners, but some research findings do not support that claim (Ross & Morrison, 1989; Williams, 1993). However, Nichols (1984) explained that learners most often select the difficulty level of a learning task "close to their perceived level of competence" (p. 333), which means learners need the ability to choose the content or path that best meet their needs if they are going to be engaged. As such, this dimension investigates if an app includes that option for learners.

Example App for Learner Control. Apps that allow learners to choose the difficulty level or setting of the content or instructional activities they will engage score well on this dimension. For example, *Subitize Tree HD* is an app designed to develop learners' perception skills by showing them a group of objects for a brief moment and then having them respond to how many items they saw. Because some learners may have more advanced subitizing skills than others, this app

allows them to adjust the length of time they are shown the objects and the amount of objects they are shown. These features then give learners the ability to set the app specifically for their ability level. Because learners control multiple settings for the learning task they are to complete, *Subitize HD* scores a 5 on this dimension.

C2. Interactivity

The “Interactivity” dimension addresses whether an app creates an engaging instructional experience for learners based on how they actively interact with the app. According to constructivist learning theory, learners create deeper understandings and knowledge of the world through experiencing the phenomena they are learning by interacting with it somehow (Duffy & Jonassen, 1992; Huba & Freed, 2000). Therefore, apps requiring learners to actively engage its content by answering question, making decisions, or performing other observable actions (Hannafin, 1992; Kearsley, 1988) score highly on this dimension.

Example App for Interactivity. Apps requiring learners to make quick decisions, such as *The Opposites*, score well on this dimension. *The Opposites* builds learners’ knowledge of vocabulary words and their commonly used antonyms. In this game-based app, words appear in bubbles and learners must correctly tap two words that are opposite of each other, which is an activity learners will likely find engaging. Because learners are making selections that represent their knowledge of the words shown in the game, *The Opposites* score a 5 on this dimension.

C3. Pace

The “Pace” dimension explores if an app allows learners to completely control the rate at which they move through the content, and this is the third new dimension being put forward. When content is only presented to learners, their motivation to engage it may wane. Yet, when learners are able to control the rate at which the content is presented to them, the added control would likely be a motivating factor. Therefore, this dimension analyzes if learners can regulate the speed at which they engage its content.

Example App for Pace. Apps that allow learners to select the type and topic of the content and how they can progress through it score best on this dimension. For example, the *Khan Academy* app contains a library of videos learners can view about an array of topics. When viewing a video in this app, learners are able to fast forward, rewind, and replay the video at their discretion. In this way, learners are able to control the rate at which they move through the video’s contents, and the *Khan Academy* app scores a 5 on this dimension.

C4. Personal preferences

The “Personal Preferences” dimension analyzes whether an app allows learners to personalize it by setting individual preferences, and this area represents another dimension that went unaddressed in any previous app evaluation rubric. According to user experience design principles, visual elements, such as color and typography, possess significant power over a learner’s perceptions in the psychological and affective domains (Garrett, 2010). The subtle changes in color or typography can evoke strong emotional reactions pertaining to joy, excitement, fear, and anger (Shneiderman & Plaisant, 2005). In Kidd’s (2005) study, visual appeal of online courses affected the learner’s level of interest and desire to learn from the course. As such, this dimension is concerned with learners being able to adjust an app’s background colors, font typography, or avatar to their liking, which encourages them to interact with the app as they pursue its learning goals.

Example App for Personal Preferences. Apps allowing learners maximum amounts of accessible customizable features that can be used when creating a learning artifact or engaging a learning activity score well on this dimension. For example, *Prezi* allows learners to create original

presentations that can be used to illustrate their conceptual understanding of a topic, and these presentations can be shared over the Internet. With *Prezi*, learners can choose color schemes, texts, images, graphics, paths, and the background of their presentation. These features then allow for each presentation created using *Prezi* to be an original experience, and it scores a 5 on this dimension.

C5. Interest

The “Interest” dimension analyzes if an app will likely appeal to its target audience. Interest is defined as the enjoyment a learner experiences when doing a task or as the enjoyment a learner experiences in the content matter of the task (Eccles & Wigfield, 1995). When learners enjoy the task they are asked to do or the content they are asked to engage, they are most likely to find the app they are using more appealing. As such, this dimension considers if the learner’s interest is matched to the app’s learning task(s) and/or content.

Example App for Interest. Apps that present their content in game formats that are easily understood by learners or include content that learners may find relevant to one of their personal interests score well on this dimension. For example, *Sushi Monster* uses a game format to build learners’ addition and multiplication skills, which are skills learners are likely to find relevant to their academic success. Because *Sushi Monster* provides an enjoyable format for young learners to review foundational mathematic skills, there is a strong likelihood that this app will be appealing to them. Because of its high potential to be appealing, *Sushi Monster* scores a 5 on this dimension.

C6. Aesthetics

The “Aesthetics” dimension evaluates an app’s graphics and interface. This dimension analyzes the artistic aspects of an app in terms of it possessing beauty and elegance (Reeves & Harmon, 1993). An app that has a visually appealing interface will likely draw learners’ attention and increase their engagement levels. In response, this dimension analyzes if an app possesses those attributes.

Example App for Aesthetics. Apps that contain dazzling images and high-quality pictures, such as *Sphere 360*, score well on this dimension. In *Sphere 360*’s library, it contains a collection of 360-degree images learners can view from all angles. The images are presented using crisp, high-definition resolutions and show unique scenes from across the globe. Given how *Sphere 360* couples its quality images with its unique interface, it scores a 5 on this dimension.

C7. Utility

The “Utility” dimension considers if learners are likely to view the skills developed and knowledge offered by an app as being important to their academic, professional, and personal lives. Utility is defined as the “usefulness of the task for individuals in terms of their future goals, including career goals...[and] is related more to the ends in the means-ends analysis of a task” (Pintrich & Schunk, 1996, p. 295). When learners can envision how the skill or knowledge they gain from an app will benefit them, they are more likely to be motivated to engage the app. As such, this dimension was designed to analyze if learners would view the app’s content as benefiting them in the future.

Example App for Utility. Apps that allow students to conduct research about topics they find relevant and can be used for academic, professional, and/or personal purposes score highly on this dimension. For example, the *Phlo – Quick Search Browsers* allows learners to research any topic using multiple search engines simultaneously. With this app, learners can quickly research topics related to their academic, professional, or personal interests. Because learners will be able to con-

nect the variety of information provided by this app to multiple facets of their life, *Phlo – Quick Search Browsers* scores a 5 on this dimension.

Considerations for Using the Rubric

To evaluate apps' educational potential effectively using this rubric, the following recommendations are put forward as supports to individuals who plan to use this rubric.

1. *Understand how instructional apps can be classified.* Individuals must be able to label apps as skills-based, content-based, or function-based depending their purpose (Cherner et al., 2014). As has been previously discussed, instructional apps are going to score differently on the rubric based on their design. Therefore, by being able to label an app into one of the three categories, individuals will be more likely to correctly assign a score of N/A when warranted.
2. *Develop a deep understanding of the rubric.* Even though the rubric is divided into three domains with multiple dimensions, it is still comprehensive. To increase their accuracy when evaluating apps, individuals may need to practice using the rubric before sharing their evaluations.
3. *Take time to become familiar with the app.* Even though it may take minimal time to classify the app based on its design and purpose, there are several nuances that can be discovered about an app. For example, it takes time to examine how students' work is saved after they exit an app and then reopen it, if and how different cultures are represented, and what supports are in place if students answer questions incorrectly. These nuances need to be found and considered before an app can be fully evaluated.
4. *Be able to separate how an app could be used in the classroom versus how it functions as a standalone tool.* Educational apps by themselves are isolated computer programs that serve a specific function. Given that teachers can create learning activities that capitalize on apps' functionalities, individuals need to be disciplined to evaluate apps as they are – being standalone computer programs – and not how they could be used in the classroom.

With these considerations in place, a narrative explaining how to use them is offered below when evaluating an app—*Google Earth*.

Google Earth is a popular iOS app. *Google Earth* provides learners with a map of the globe, and it allows them to input and search for addresses, names of landmarks, and geographical features. When one of those criteria is inputted, *Google Earth* will take learners to that location, and they can manipulate their view of that location using their fingers. Plus, learners can research a location by connecting to Wikipedia (www.wikipedia.com) and by connecting to a linked website (if a website is offered). Learners are also given the option to turn on or off map features, such as places, businesses, borders, and roads among others. By spending time with *Google Earth* to discover these features, evaluators will have also come to understand that *Google Earth* is a content-based app because it presents map views to learners but does not require them to answer questions or create learning artifacts. Once evaluators come to these understandings of *Google Earth's* functionalities and options, they are able to use the rubric to evaluate the app. To do so, evaluators should start with Domain A and analyze *Google Earth* using its indicators before progressing to domains B and C. Table 2 provides scores for *Google Earth* based on the rubric.

Table 2: Google Earth Evaluation

Domain A: Instruction	Domain B: Design	Domain C: Engagement
A1. Rigor: 2	B1. Ability to Save Progress: 1	C1. Learner Control: N/A
A2. 21 st Century Skills: 2	B2. Integration: 5	C2. Interactivity: 3
A3. Connections to Future Learning: 3	B3. Screen Design: 4	C3. Pace: 5
A4. Value of Errors: N/A	B4. Ease of Use: 4	C4. Personal Preferences: 3
A5. Feedback to Teacher: N/A	B5. Navigation: 4	C5. Interest: 5
A6. Level of Learning Material: 5	B6. Goal Orientation: 5	C6. Aesthetics: 5
A7. Cooperative Learning: 1	B7. Information Presentation: 5	C7. Utility: 5
A8. Accommodation of Individual Differences: 1	B8. Media Integration: 5	
	B9. Cultural Sensitivity: 4	

Limitations and Future Work

Although this rubric was designed to be as comprehensive and inclusive as possible, there are limitations for its use. In this context, limitations are defined as the constraints or factors that impede the rubric's generalizability (Given, 2008), and these limitations affect how the rubric can be used.

First, because a variety of apps exists and each variety has specific purposes (Cherner et al., 2014), creating a single rubric to evaluate all varieties of educational apps is not possible. For example, the criteria used to evaluate an instructional app that is designed to support student learning as described above will be fundamentally different than the criteria used to evaluate a teacher resource app that is designed to support educators in planning instruction and evaluating student learning. Even though these two varieties of apps are designed for education, their purposes are unique from each other. As such, creating a comprehensive rubric for teacher resource apps is still needed.

Furthermore, although assessing instructional apps was the purpose for creating this rubric, app evaluators must understand that an app's design must be considered when deciding to score an app versus assigning it an N/A. For example, when considering the *Feedback to Teacher* dimension, app evaluators are analyzing how teachers can monitor student progress. Therefore, when evaluating a content-based app that only provides information to learners (Cherner et al., 2014), the app should be scored as a N/A because it was designed to only present content to learners, not to assess their learning or use of that content. However, skill-based and function-based app should be scored using a numerical indicator because learners will engage those app's content by answering questions, creating learning artifacts, or engaging it in some other manner (Cherner et al., 2014). As such, app evaluators must consider an app's design and purpose when assessing it using this rubric.

Second, while this rubric was designed with the intention to quantify different aspects of instructional apps, the act of interpreting apps still represents a qualitative activity because it requires the evaluator to be used as an evaluative instrument (Patton, 2002; Strauss & Corbin, 1990). This means evaluators must interpret the rubric, apply it to the app being evaluated, and make deci-

sions about how multiple components of an app score on the rubric based on their understanding of the rubric, the app, and the evaluation process. Even though the purpose of rubrics is to objectively evaluate a phenomenon, an interpretive, subjective element is still inherent in that process.

Third, the practical use of this rubric by classroom teachers is questionable. Although this rubric provides a tool for evaluating apps comprehensively, it would require teachers to spend a significant amount of time learning how to use it appropriately, and that is before considering the time they would need to evaluate apps. As such, this rubric may be more meaningful for educational app resources to use, and examples of these resources include the materials posted by Kathy Schrock (n.d.) (<http://www.kathyschrock.net/>), Appitic (Sada, 2012) (www.appitic.com), and Common Sense Media's Graphite (2014) (www.graphite.org). Currently, App Ed Review (2013) (www.appedreview.org) is the only resource using this rubric to comprehensively evaluate apps. Additionally, app designers will find value in this rubric, as they will be able to design their app according to its research-based dimensions. As such, this rubric will be important for them to consult as they make design and functionality decisions when planning their app.

Fourth, the worth of an educational app is determined by how it can be effectively integrated into classroom teaching. Even if an app earns high scores on this rubric's multiple dimensions, teachers must still integrate that app meaningfully into their instruction by aligning it to their content and pedagogy for it to be used effectively (Koehler & Mishra; 2009). An app in isolation is not a high-quality educational tool; however, an app integrated effectively and in conjunction with sound teaching practices in a way that supports student learning is a high-quality educational tool.

Fifth, as educational technologies continue to emerge, future researchers must review and rework the different dimensions of this rubric. Whereas the rubrics designed to evaluate the educational software of the 1990s have become outdated due to the rise of tablet technologies, this rubric will eventually become outdated with the advent of new technologies in the future. In this way, this rubric is not a static artifact, but it is one that will offer future researchers a starting point for developing a new rubric aligned to those yet invented technologies. For example, Google Glass, a form of eyewear that changes how wearers interact with their reality, represents a new generation of hardware that has potential to change how individuals interface with the world (Tsukayama, 2012), similar to how mobile computing devices are currently changing education (Dundar & Akcayir, 2012; Waters, 2010). Although Google Glass has not been widely adopted yet by the general public, it is foreseeable that it, or a similar type of emerging technology, will become available and mainstreamed into society. As that happens, the field of education will be affected, and schools will likely begin integrating this new technology into the classroom. With the integration of new educational technologies, improvements and modifications to this rubric will need to be made.

Conclusion

In order for teachers to select quality apps to use with their students, they need the quality resources to do so. The evaluation rubric introduced in this article is a response to that need. This rubric was designed using previously published research while also taking into consideration the changes in education since those works were published. As emerging technologies continue to be developed and integrated into schools and classroom across the United States and globe, this rubric is crucial instrument because it offers researchers, teachers, and app developers a common metric for evaluating the quality of educational apps. Without this common set of criteria, making definitive statements about the quality of educational apps can quickly become a guessing game.

References

- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Allyn & Bacon.
- App Ed Review. (2013). *Home*. Retrieved from www.appedreview.org
- Assessment & Teaching of 21st Century Skills (2014). *What are 21st-Century Skills? Assessment & Teaching of 21st Century Skills*. Retrieved May 15, 2014 from <http://atc21s.org/index.php/about/what-are-21st-century-skills/>
- Berge, Z. L. (1995). The role of the online instructor/facilitator. *Educational Technology*, 35(1), 22-30.
- Blackburn, B. R. (2014). *Rigor in your classroom: A toolkit for teachers*. Routledge.
- Blumenfeld, P. C., Marx, R. W., Soloway, E., & Krajcik, J. (1996). Learning with peers: From small group cooperation to collaborative communities. *Educational Researcher*, 37-40.
- Buckler, T. (2012). Is there an app for that? Developing an evaluation rubric for apps for use with adults with special needs? *The Journal of BSN Honors Research*, 5(1), 19-32.
- Cherner, T., Dix, J., & Lee, C. (2014). Cleaning up that mess: A framework for classifying educational apps. *Contemporary Issues in Technology and Teacher Education*, 14(2). Retrieved from <http://www.citejournal.org/vol14/iss2/general/article1.cfm>
- Collins, A. (1991). The role of computer technology in restructuring schools. *Phi Delta Kappan*, 73(1), 28-36.
- Common Sense Media. (2014). *Graphite*. Retrieved from <https://www.graphite.org/>
- Coughlan, J., & Morar, S. S. (2008). Development of a tool for evaluating multimedia for surgical education. *Journal of Surgical Research*, 149(1), 94-100.
- Cronjé, J. (2006). Paradigms regained: Toward integrating objectivism and constructivism in instructional design and the learning sciences. *Educational Technology Research and Development*, 54(4), 387-416.
- Csikszentmihalyi, M. (1997). *Flow and the psychology of discovery and invention*. New York: HarperPerennial.
- Daggett, W. (2005, September). *Achieving academic excellence through rigor and relevance*. International Center for Leadership in Education. Retrieved May 15, 2014 from <http://www.dupage.k12.il.us/pdf/121306%20Achieving%20Academic%20Excellence%20through%20Rigor%20and%20Relevance.pdf>
- Duffy, T. M., & Jonassen, D. H. (Eds.). (1992). *Constructivism and the technology of instruction: A conversation*. Psychology Press.
- Dundar, H., & Akcayir, M. (2012). Tablet vs. paper: The effect on learners' reading performance. *International Electronic Journal of Elementary Education*, 4(3), 441-450.
- Earl, M. (2013). iPads in the social classroom, K-12 education. *Instructional Technology Education Special Research papers*, Paper 5.
- Eccles, J., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, 21, 215-225.
- Ehlers, U. D., & Pawlowski, J. M. (2006). *Handbook on quality and standardisation in E-learning*. Heidelberg: Springer.
- Elawar, M. C., & Corno, L. (1985). A factorial experiment in teachers' written feedback on student homework: Changing teacher behavior a little rather than a lot. *Journal of educational psychology*, 77(2), 162-173.
- Elissavet, G., & Economides, A. A. (2003). An evaluation instrument for hypermedia courseware. *Educational Technology & Society*, 6(2), 31-44.

A Comprehensive Evaluation Rubric

- Fisher, D., & Frey, N. (2013). *Common core English language arts in a PLC at work, grades 3-5 [electronic resource]* / Douglas Fisher, Nancy Frey; foreword by Rebecca DuFour. Bloomington, Indiana: Solution Tree Press, c2013.
- Galitz, W. O. (1985). *Handbook of screen format design* (2nd ed.), Wellesley Hills, MA: QED Information Sciences.
- Galitz, W. O. (1992). *User-interface screen design*. John Wiley & Sons.
- Garrett, J. J. (2010). *The elements of user experience: User-centered design for the web and beyond*. Pearson Education.
- Given, L. (2008). *The SAGE encyclopedia of qualitative research methods*. Thousand Oaks CA: Sage.
- Gokhale, A. A. (1995). Collaborative learning enhances critical thinking. *Journal of Technology Education*, 7(1), 22-30.
- Graham, C. R. (2005) Blended learning systems: Definitions, current trends and future directions. In C. J. Bonk & C. R. Graham (Eds.), *Handbook of blended learning: Global Perspectives, local designs* (pp. 3-21). San Francisco, CA: Pfeiffer Publishing.
- Greenfield, A. (2006). *Everyware: The dawning age of ubiquitous computing*. Berkeley, CA: New Riders.
- Hannafin, M. J. (1992). Emerging technologies, ISD, and learning environments: Critical perspectives. *Educational Technology Research and Development*, 40(1), 49-63.
- Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238.
- Herman, J., & Linn, R. (2013). *On the road to assessing deeper learning: The status of smarter balanced and PARCC assessment consortia. CRESST Report 823*. National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Herrington, J., & Oliver, R. (1997). Multimedia, magic and the way students respond to a situated learning environment. *Australian Journal of Educational Technology*, 13, 127-143.
- Hess, K. K., Carlock, D., Jones, B., & Walkup, J. R. (2009). *What exactly do "fewer, clearer, and higher standards" really look like in the classroom? Using a cognitive rigor matrix to analyze curriculum, plan lessons, and implement assessments*. Retrieved May 15, 2014 from http://www.nciea.org/publications/cognitiverigorpaper_KH12.pdf
- Huba, M.E. & Freed, J.E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning*. Needham Heights, MA: Allyn & Bacon.
- The Institute of Museum and Library Services. (2014). Museums, Libraries, and 21st Century Skills. About Us. Retrieved May 15, 2014 from http://www.imls.gov/about/21st_century_skills_list.aspx
- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38(12), 1217-1218.
- Jeffries, P. R. (2005). A frame work for designing, implementing, and evaluating simulations used as teaching strategies in nursing. *Nursing Education Perspectives*, 26(2), 96-103.
- Johnson, D. W., & Johnson, R. T. (1987). *Learning together and alone: Cooperative, competitive, and individualistic learning*. Prentice-Hall
- Jonassen, D. H., & Grabowski, B. L. (2012). *Handbook of individual differences learning and instruction*. Routledge.
- Juul, J. (2004). Introduction to game time/time to play – An examination of game temporality. In N. Wadrip-Fruin & P. Harrigan (Eds.), *First Person: New Media as Story, Performance and Game* (pp. 131-142). Cambridge, MIT Press.
- Kearsley, G. (1988). *Online help systems: Design and implementation*. Norwood, NJ: Alex Publications.
- Kennedy, G., Petrovic, T., & Keppell, M. (1998). The development of multimedia evaluation criteria and a program of evaluation for computer aided learning. *ASCILITE*, 98, 407-415.

- Kidd, T. (2005). Key aspects affecting students' perception regarding the instructional quality of online and web based courses. *International Journal of Instructional Technology and Distance Learning*, 2(10), 55-61.
- Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary Issues in Technology and Teacher Education*, 9(1), 60-70.
- Kulik, C. L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, 7(1), 75-94.
- Leung, C. H., & Chan, Y. Y. (2003, July). Mobile learning: A new paradigm in electronic learning. In *Advanced Learning Technologies, 2003. Proceedings. The 3rd IEEE International Conference* (pp. 76-80). IEEE.
- Lucas, L. (1991). Visually designing the computer-student interface. *Educational Technology*, 31(7), 56-58.
- Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing research*, 35(6), 382-386.
- Matell, M. S., & Jacoby, J. (1971). Is there an optimal number of alternatives for Likert scale items? I. Reliability and validity. *Educational and Psychological Measurement*, 31(3), 657-674.
- Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83, 484-490.
- Mayer, R. E., & Anderson, R. B. (1992). The instructive animation: Helping students build connections between words and pictures in multimedia learning. *Journal of Educational Psychology*, 84, 444-452.
- Mayer, R. E., & Moreno, R. (1998). A cognitive theory of multimedia learning: Implications for design principles. In N. H. Naryanan (Ed.), *Electronic proceedings of the CHI'98 workshop on hyped-media to hyper-media: Toward theoretical foundations of design, use and evaluation*.
- Murray, O. T., & Olcese, N. R. (2011). Teaching and learning with iPads, ready or not? *TechTrends*, 55(6), 42-48.
- National Center on Education and the Economy. (2013). *What does it really mean to be college and work ready?* Washington DC: National Center on Education and the Economy.
- National Council for the Social Studies. (2013). *The college, career, and civic life (C3) framework for social studies state standards: Guidance for enhancing the rigor of K-12 civics, economics, geography, and history*. Silver Spring, MD: National Council for Social Studies.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards*. Washington, DC: National Governors Association Center for Best Practices and the Council of State School Officers.
- Next Generation Science Standards Lead States. (2013). *Next Generation Science Standards: For States, By States*. Washington DC: The National Academies Press.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review*, 91(3), 328-346.
- Ovando, M. N. (1994). Constructive feedback: A key to successful teaching and learning. *International Journal of Educational Management*, 8(6), 19-22.
- Overbaugh, R. C. (1994). Research-based guidelines for computer-based instruction development. *Journal of Research on Computing in Education*, 27(1), 29-47.
- Palloff, R. M., & Pratt, K. (1999). *Building learning communities in cyberspace*. San Francisco: Jossey-Bass Publishers.
- Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. New York, New York: Basic Books.

A Comprehensive Evaluation Rubric

- Partnership for 21st Century Skills. (2009). *Curriculum and instruction: A 21st century skills implementation guide*. Tucson, Arizona: Partnership for 21st Century Skills. Retrieved May 15, 2014 from http://www.p21.org/storage/documents/p21-stateimp_curriculuminstruction.pdf
- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Phillips, R. (2005). Challenging the primacy of lectures: The dissonance between theory and practice in university teaching. *Journal of University Teaching & Learning Practice*, 2(1), 2.
- Picciano, A. G., & Seaman, J. (2007). K-12 online learning: A survey of US school district administrators. *Journal of Asynchronous Learning Networks*, 11(3).
- Pilgrim, J., Bledsoe, C., & Reily, S. (2012). New technologies in the classroom. *Delta Kappa Gamma Bulletin*, 78(4).
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667.
- Pintrich, P. R., & Schunk, D. H. (1996). *Motivation in education: Theory, research and applications*. Englewood Cliffs: Prentice Hall.
- Powell, G. C. (1993). *Incorporating student cultural diversity into instructional systems design: An investigation of faculty awareness and teaching practices*. (Doctoral dissertation, University of Georgia).
- Premkumar, K., Hunter, W., Davison, J., & Jennett, P. (1998). Development and validation of an evaluation tool for multimedia resources in health education. *International Journal of Medical Informatics*, 50(1), 243-250.
- Rao, L. (2012, January 19). Apple: 20,000 education iPad apps developed; 1.5 million devices in use at schools. *TechCrunch*. Retrieved May 13, 2014, from <http://techcrunch.com/2012/01/19/apple-20000-education-ipad-apps-developed-1-5-million-devices-in-use-at-schools>
- Reeves, T. C. (1994). Evaluating what really matters in computer-based education. *Computer Education: New Perspectives*, 219-246.
- Reeves, T. C., & Harmon, S. W. (1993). Systematic evaluation procedures for interactive multimedia for education and training. *Multimedia Computing: Preparing for the 21st Century*, 472-505.
- Reeves, T. C., & Hedberg, J. G. (2003). *Interactive learning systems evaluation*. Englewood, NJ: Educational Technology Publications.
- Reeves, T. C., & Reeves, P. M. (1997). Effective dimensions of interactive learning on the World Wide Web. *Khan*, 62, 59-66.
- Rieber, L. P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development*, 44(2), 43-58.
- Ross, S. M., & Morrison, G. R. (1989). In search of a happy medium in instructional technology research: Issues concerning external validity, media replications, and student control. *Educational Technology Research and Development*, 37(1) 19-33.
- Sada, A. (2012). *App lists for education*. Retrieved from <http://www.appitic.com/>
- Schibeci, R., Lake, D., Phillips, R., Lowe, K., Cummings, R., & Miller, E. (2008). Evaluating the use of learning objects in Australian and New Zealand schools. *Computers & Education*, 50(1), 271-283.
- Schrock, K. (n.d.). *Home*. Retrieved from <http://www.kathyschrock.net/>
- Sherr, S. (1979). *Electronic displays*. New York: Wiley-Interscience.
- Shiratudin, N., & Landoni, M. (2002). Evaluation of content activities in children's educational software. *Evaluation and Program Planning*, 25(2), 175-182.
- Shneiderman, B., & Plaisant, C. (2005). *Designing the user interface: strategies for effective human-computer interaction*. Boston: Pearson/Addison Wesley.

- Slavin, R. E. (1990a). *Cooperative learning: Theory, research, and practice*. Englewood Cliffs, NJ: Prentice Hall.
- Slavin, R. E. (1990b). IBM's writing to read: Is it right for reading? *Phi Delta Kappan*, 72, 214-216.
- Slavin, R. E. (1992). Cooperative learning. In M. C. Alkin (Ed.), *Encyclopedia of educational research* (pp. 235-238). New York: Macmillan.
- Squires, D., & Preece, J. (1999). Predicting quality in educational software: Evaluating for learning, usability and the synergy between them. *Interacting with Computers*, 11(5), 467-483.
- Strauss, A. C. & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage Publications.
- Tomlinson, C. (1999). *The differentiated classroom: Responding to the needs of all learners*. Alexandria, VA: ASCD.
- Tomlinson, C. (2001). *How to differentiate instruction to mixed-ability classrooms* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Trochim, W. (2006). Likert scaling. *Web Center for Social Research Methods*. Research Methods Knowledge Base. Retrieved May 13, 2014, from <http://www.socialresearchmethods.net/kb/scallik.php>
- Tsukayama, H. (2012). Everything you need to know about Google Glass. *The Washington Post*. Retrieved from <http://www.washingtonpost.com/blogs/the-switch/wp/2014/02/27/everything-you-need-to-know-about-google-glass/>
- Vygotsky, L. (1978). Interaction between learning and development. In M. Gauvain & M. Cole (Eds.) *Readings on the development of children* (pp. 34-41).
- Walker, H. (2010). Evaluating the effectiveness of apps for mobile devices. *Journal of Special Education Technology*, 26(4), 59-66.
- Waters, J. K. (2010). Enter the iPad (or Not?). *THE Journal*, 37(6), 38-40.
- Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education*. Madison, WI: National Institute for Science Education.
- Webb, N. L. (1999). Alignment of science and mathematics standards and assessments in four states. Unpublished manuscript. (ERIC Document Service No. ED440852).
- Williams, M. D. (1993). A comprehensive review of student-control: The role of student characteristics. In M. R. Simonson (Ed.), *Proceedings of the Annual Conference of the Association for Educational Communications and Technology* (pp. 1083-1114). New Orleans, LA: Association for Educational Communications and Technology.
- Winslow, J., Dickerson, J., & Lee, C. (2013). *Applied Technologies for Teachers*. Dubuque, IA: Kendall Hunt.

APPENDIX A

Evaluation Rubric for Educational Apps

A. Instruction: The following dimensions are used to measure an app's educational value.					
A1. Rigor: Which level of Webb's Depth of Knowledge must students engage to meet the app's learning objective?					
5	4	3	2	1	NA
The app's content best aligns to Level Four "Extended Thinking" skills.	The app's content best aligns to Level Three "Strategic Thinking" skills.	The app's content best aligns to Level Two "Skill/Concept" thinking skills.	The app's content best aligns to Level One "Recall Reproduction" thinking skills.	The app does not align to Webb's Depth of Knowledge because it is geared more toward entertaining students than educating them.	Not Applicable
A2. 21 st Century Skills: Does the app require users to engage "21 st Century" skills, which includes the ability to collaborate, make data-driven decisions, and solve complex problems?					
5	4	3	2	1	NA
The app frequently requires users to engage 21 st Century skills.	The app sometimes requires users to engage 21 st Century skills.	The app rarely requires users to engage 21 st Century skills.	The app does not require users to engage 21 st Century skills.	The app promotes rote memorization skills.	Not Applicable
A3. Connections to Future Learning: Does the app's content build users' literacy or numeracy skills so they are more prepared to engage future content area learning and are on track to become "college and career" ready?					
5	4	3	2	1	NA
The app's content targets and teaches specific literacy or numeracy skills to better prepare users for future content area learning, which promotes "college and career" readiness.	The app's content targets and teaches general literacy or numeracy skills to better prepare users for future content area learning, which promotes "college and career" readiness.	The app's content targets and reinforces specific literacy or numeracy skills to better prepare users for future content area learning, which aligns to "college and career" readiness.	The app's content targets and reinforces general literacy or numeracy skills to better prepare users for future content area learning, which aligns to "college and career" readiness.	The app's content does not connect literacy and numeracy skills with future content area learning and "college and career" readiness is not promoted.	Not Applicable

A4. Value of Errors: If users answer a question incorrectly, does the app use the incorrect answer to advance learning?					
5	4	3	2	1	NA
The app offers users an explanation for their incorrect answer, allows users to review the material related to the question, and then gives them an opportunity to answer the question again.	The app offers users either an explanation for their incorrect answer or directs them to the material related to the question before giving users an opportunity to answer the question again.	The app offers users an explanation for why their answer was incorrect, but the app does not give them an opportunity to answer the question again.	The app offers users an opportunity to answer the question again, but the app does not provide information to explain why their answer was incorrect the first time.	The app does not give users an opportunity to answer the question again or provide information about why their answer was incorrect.	Not Applicable
A5. Feedback to Teacher: How does the app allow teachers to monitor their students' progress?					
5	4	3	2	1	NA
The app tracks students' academic growth, achievements, and progress made while using the app. Teachers can access the reports either through email or by visiting the app's website.	The app tracks students' achievements or progress made when using the app, but it does not track students' growth. Teachers can access the reports using email or by visiting the app's website.	The app tracks students' achievements or progress made when using the app, but it does not track students' growth. Also, student progress can only be seen in the app and not by using email or a website.	The app only tracks students' achievements or progress when they are using the app. Once the app is exited, progress is not saved, and teachers do not have any way to access it.	The app does not track students' academic growth, achievements, or progress made.	Not Applicable
A6. Level of Material: Is the app's content for specific grade-levels appropriate for its target audience?					
5	4	3	2	1	NA
The app's content for specific grade-levels is always matched appropriately for its target audience.	The app's content for specific grade-levels is usually matched appropriately for its target audience.	The app's content for specific grade-levels is rarely matched appropriately for its target audience.	The app's content for specific grade-levels is not matched appropriately for its target audience.	The app's content is inappropriate for educational purposes.	Not Applicable

A7. Cooperative Learning: Does the app allow users to communicate (e.g. send/receive messages and post to message boards), share learning artifacts (e.g. images, recordings, and documents), and collaborate (e.g. create presentations and larger projects) with other users?					
5	4	3	2	1	NA
The app features an online community where users can communicate freely with other users, share learning artifacts, and collaborate on projects.	The app features an online community where users may be able to communicate with other users, share learning artifacts, or collaborate on projects.	The app features an online community where users can communicate freely with other users, but they are not able to share learning artifacts or collaborate on projects.	The app includes an online community, but users have little or no control over how they communicate with other users. Also, the ability to share learning artifacts or collaborate on projects is not available.	The app does not allow users to communicate or interact with other users.	Not Applicable
A8. Accommodation of Individual Differences: Does the presentation of the app's content recognize that users learn in uniquely different ways (e.g. visual, auditory, and/or kinesthetic)?					
5	4	3	2	1	NA
All of the app's content recognizes that users learn in different ways and caters to differentiated instructional practices.	Most of the app's content recognizes that users learn in different ways and caters to differentiated instructional practices.	Some of the app's content recognizes that users learn in different ways and caters to differentiated instructional practices.	Very little of the app's content recognizes that users learn in different ways and caters to differentiated instructional practices.	The app's content is presented in a format that assumes all users learn in the same way.	Not Applicable

B. Design: The following dimensions are used to evaluate an app's overall functionality.					
B1: Ability to Save Progress: Does the app allow users to return to the content they were last engaging after exiting the app?					
5	4	3	2	1	NA
After exiting an app, users can reopen the app and automatically return to the content they were last engaging when they logged off.	After exiting an app, users can reopen the app and resume engaging it automatically in close proximity to where they were when they logged off.	After exiting an app, users can reopen the app and manually select the content they were last engaging when they logged off.	After exiting an app, users can reopen it and manually select the content they were last engaging when they logged off, but the content may be different.	After exiting an app, users must begin on the first level when returning to the app.	Not Applicable
B2. Integration: Is the app enhanced by how it connects to (1) other apps, (2) online communities, (3) independent websites, and (4) users' email?					
5	4	3	2	1	NA
The app is enhanced with how it integrates with all four of the listed connections.	The app is enhanced with how it integrates with three of the listed connections.	The app is enhanced with how it integrates with two of the listed connections.	The app is enhanced with how it integrates with only one of the listed connections.	The app does not integrate with any of the listed connections.	Not Applicable
B3. Screen Design: Is the app's text, graphics, videos, sound, and speech well-organized?					
5	4	3	2	1	NA
The app's text, graphics, videos, sound, and speech are always organized in a way that enhances the app's content.	The app's text, graphics, videos, sound, and speech are usually organized in a way that complements the app's content.	The app's text, graphics, videos, sound, and speech are organized in a way that does not enhance or detract from the app's content.	The app's text, graphics, videos, sound, and speech are not well organized and may detract from the app's content.	The app's text, graphics, videos, sound, and speech are cluttered and confusing, which detracts from the app's content.	Not Applicable

B4. Ease of Use: Is the app intuitive and are users able to engage it with minimal guidance?					
5	4	3	2	1	NA
Users are able to engage the app immediately with no guidance.	Users are able to engage the app immediately with minimal guidance.	Users are able to engage the app, but only after some guidance.	Users are able to engage the app, but only after substantial guidance.	Users are able to engage the app, but training materials are continually needed to do so.	Not Applicable
B5. Navigation: How easily can users move through the app's content and options?					
5	4	3	2	1	NA
Users can move through the app's content and options fluidly.	Users need to put forth some effort to move through the app's content and options.	Users need to make multiple clicks and/or swipes to move through the app's content and options.	Users are somewhat impeded from moving fluidly through the app's content and options because of its organization.	Users encounter substantial challenges when trying to move through the app's content and options because of its disjointed organization.	Not Applicable
B6. Goal Orientation: Does each component of the app contribute to users learning the intended objective?					
5	4	3	2	1	NA
All components of the app contribute to users learning the objective.	Most components of the app contribute to users learning the objective.	Some components of the app contribute to users learning the objective.	Few components of the app contribute to users learning the objective.	None of the app's components contribute to users learning the objective.	Not Applicable

B7. Information Presentation: Is the app's content presented in a logical manner? (e.g. the app's content grows increasingly rigorous as users experience success, the app activates users' background knowledge before presenting them new information, and/or the app provides an overview of its content before users engage specific tutorials or activities.)					
5	4	3	2	1	NA
The app's content is presented in a logical manner.	The app's content is presented in a manner that is mostly logical.	The app's content is presented in a manner that is somewhat logical.	The app's content is presented in a manner that is somewhat illogical.	The app's content is presented in an illogical manner.	Not Applicable
B8. Media Integration: Are the app's texts, graphics, videos, sounds, and speech integrated effectively so each of the app's media components complements each other and forms a cohesive program?					
5	4	3	2	1	NA
The app's text, graphics, videos, sounds, and speech are integrated seamlessly to form a cohesive program.	The app's text, graphics, videos, sounds, and speech are integrated adequately to form a cohesive program.	The app's text, graphics, videos, sounds, and speech are integrated poorly, but they still form a mostly cohesive program.	The app's text, graphics, videos, sound, and speech are mismatched to the app's content, but they still form a somewhat cohesive program.	The app's text, graphics, videos, sound, and speech are mismatched to the app's content, and it does not form a cohesive program.	Not Applicable
B9. Cultural Sensitivity: Does the app use culturally responsive teaching methods to represent diverse populations?					
5	4	3	2	1	NA
The app allows users to connect to and share ideas with people from diverse communities across the globe.	The app allows users to explore diverse cultures from across the globe using authentic pictures, images, and/or texts.	The app makes general references to one or more cultures.	The app makes reference to diverse cultures, but the information contains some cultural biases.	The app presents stereotypical images and information about cultures.	Not Applicable

C. Engagement: The following dimensions are used to evaluate an app's motivational potential.						
C1. Learner Control: Does the app allow users to select the level (e.g. grade level, age level, or difficulty level) where they will begin engaging content?						
5	4	3	2	1	NA	
The app allows users complete control related to the level at which they begin engaging its content.	The app allows users significant control related to the level at which they begin engaging its content.	The app allows users some control related to the level at which they begin engaging its content.	The app allows users little control related to the level at which they begin engaging its content.	The app allows users no control related to the level at which they begin engaging its content.		Not Applicable
C2. Interactivity: Does the app only present content to users, or does the app allow users to make decisions, answer questions, or actively engage it in some other manner?						
5	4	3	2	1	NA	
The app frequently requires users to engage its content by making decisions, answering questions, or actively engaging it in some other manner.	The app often requires users to engage its content by making decisions, answering questions, or actively engaging it in some other manner.	The app occasionally requires users to engage its content by making decisions, answering questions, or actively engaging it in some other manner.	The app only allows users to interact with it by searching for specific content (e.g. looking up a word in a dictionary app or researching terms using a search engine app).	The app only presents content to users. Users are never asked to make decisions, answer questions, or actively engage it in some other manner.		Not Applicable
C3. Pace: Does the app allow users to control the rate at which they are presented content? If the app does not allow users to control the rate at which they move through the content, does the app present the content at an appropriate rate for its target audience?						
5	4	3	2	1	NA	
The app allows users to completely control the rate at which they move through the app's content.	The app allows users to mostly control the rate at which they move through the app's content.	The app does not allow users to control the rate at which they move through the app's content, but the content is presented at an appropriate rate for its target audience.	The app does not allow users to control the rate at which they move through the app's content, but the content is presented at a mostly appropriate rate for its target audience.	The app does not allow users to control the rate at which they move through the app's content, and the content is presented at an inappropriate rate for its target audience.		Not Applicable

C4. Flexibility: Can users personalize the app by setting individual preferences (e.g. background music, images, avatars) easily?					
5	4	3	2	1	NA
Users can personalize four or more features of the app.	Users can personalize three features of the app.	Users can personalize two features of the app.	Users can personalize one feature of the app.	Users cannot personalize the app.	Not Applicable
C5. Interest: Will the app's content likely appeal to its targeted audience?					
5	4	3	2	1	NA
The app's content will most likely appeal to its targeted audience.	The app's content will likely appeal to its targeted audience.	The app's content might appeal to its targeted audience.	The app's content might not appeal to its targeted audience.	The app's content most likely will not appeal to its targeted audience.	Not Applicable
C6. Aesthetics: Will the app's graphics and interface likely motivate users to engage it?					
5	4	3	2	1	NA
The app's graphics and interface will most likely motivate users to engage it.	The app's graphics and interface may motivate users to engage it.	The app's graphics and interface may possibly motivate users to engage it.	The app's graphics and interface will likely have no impact on whether or not users are motivated to engage it.	The app's graphics and interface will likely deter users from engaging it.	Not Applicable
C7. Utility: Will users be motivated to engage the app because they see it as being valuable to their academic, professional, and personal lives?					
5	4	3	2	1	NA
Users will most likely connect the app's content to their academic, professional, and/or personal lives. This connection will result in them being motivated to use the app.	Users are likely to connect the app's content to their academic, professional, and/or personal lives. This connection may result in them being motivated to use the app.	Users may connect the app's content to their academic, professional, and/or personal lives, but that connection will not likely result in them being motivated to use the app.	Users will most likely not be able to connect the app's content to their academic, professional, and/or personal lives, and they will likely not be motivated to use the app.	Users will most likely see no value in the app's content over the course of their academic, professional, and/or personal lives, and they will likely not be motivated to use the app.	Not Applicable

APPENDIX B

Validation of the Evaluation Rubric for Tablet Apps

Thank you for your willingness to critique our instrument. Your feedback will provide us with information about how we can better refine our instrument, and please provide your feedback using the following form.

Name: Title:

Number of years that you have been an educator:

Your field of expertise:

Directions: Please review each of the following dimensions to determine how “concise” and “adequate” they are by rating them on a 5-point scale using the following breakdown:

5 = Very Good 4 = Good 3 = Average 2 = Below Average 1 = Poor

Additionally, the “concise” rating focuses on how succinct and crisp the language used to describe the dimension is. The “adequate” rating determines if the language used satisfactorily describes the dimension’s focus. Lastly, please provide any additional comments in the space following each dimension, especially if a rating of 1, 2, or 3 is assigned.

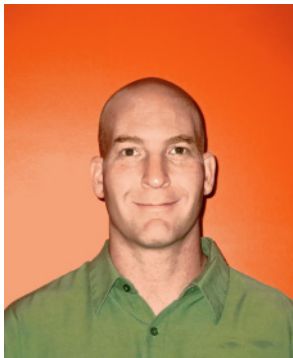
How concise is this dimension’s question ?	5	4	3	2	1
How adequate is this dimension’s question ?	5	4	3	2	1
How concise are this dimension’s indicators ?	5	4	3	2	1
How adequate is this dimension’s indicators ?	5	4	3	2	1

Additional Comment(s) and Suggestion(s):

Biographies



Cheng-Yuan “Corey” Lee earned a Ph.D. in curriculum and instruction with an emphasis on instructional system design from the University of Central Florida. Throughout his career of teaching, he developed strong interests in the areas of distance education, mobile learning, interactive multimedia, and graphic/website design. In his classes, he introduces a wide range of software application and open-source packages to students, allowing them to master computer-based tools used in the design and creation of electronic media, such as electronic images, sounds, videos, text, and motion.



Todd Sloan Cherner is an assistant professor of education at Coastal Carolina University. He specializes in preparing future high school teachers for the classroom, and he embeds instructional technology, disciplinary literacy, and Critical Theory into his teaching. Todd deeply believes that preparing PreK-12 students to become effective citizens in a global community is a multi-faceted endeavor, which involves cooperation and collaboration among many of society’s stakeholders.