Literature Review of Educational App Evaluation Rubrics

Faizal Yamimi Mustaffa, Abdul Rahim Salam, Faruk Muhammad, Ghazali Bunari and Lokman Hafiz Asary
Language Academy, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

Abstract. Statistics have shown that in the first quarter of 2018, there were about 2 million apps available for download for IOS devices while there were about 3.8 million Android apps which were available for download (Statista, 2018a). Out of these millions of apps, tens of thousands of them are educational apps. Some of these apps can help both educators and students in their teaching and learning activities. However, because of the sheer number of apps available to be downloaded for free in the market, it has become hard for educators and students to choose which ones are right for them to fulfil their needs. In order to help them determine the overall quality of these educational apps and gauge their specific strengths and weaknesses, several educational app evaluation rubrics have been proposed by researchers. This study aims to analyse these evaluation rubrics for their comprehensiveness and usefulness to users, especially for educators who want to choose applications which could complement their formal or informal teaching activities.

Keywords: Mobile assisted language learning (MALL), App evaluation, Educational app evaluation rubrics

1 Introduction

Statistics have shown that in the first quarter of 2018, there were about 2 million apps available for download for IOS devices while there were about 3.8 million Android apps which were available for download (Statista, 2018a). In early 2018, there are an estimated 25 million smartphone users in Malaysia (Statista, 2018b). This means that over 60% of Malaysians own smartphones which can make use of the millions of apps available on the IOS and Android platforms. Out of these millions of apps, tens of thousands of them are educational apps. While this, on the surface, seems to be good news to app users, it has posed several problems to them. Firstly, because there are so many apps in the market, it has made it hard for users to choose the right ones to install (Ok, Kim, Kam, & Bryant, 2016; Lee & Cherney, 2015). Furthermore, in the case of educators, many of them do not have enough time to go through the selection process without any tools to help them make the right decision. In today’s scenario, it is a very challenging task for an educator to choose the right app without any form of guidelines or rubrics to assist them in their decision-making process (Ok et al., 2016). Therefore, this study was conducted in order to find research-informed rubrics which could be used by educators to help them in choosing the right app out of the tens of thousands which are currently available in the market.
2 Methodology

In order to find the research-informed rubrics, we did a search in three academic search engines which are Google Scholar, Web of Science, and Scopus. We searched for only one key phrase which was ‘app evaluation rubrics’. The reason we used only one key phrase was because it was very specific and we believed it would generate very specific results – only researches which had anything to do with evaluation rubrics for apps. Through our search, we received less than 10 relevant results from Google Scholar, 7 results from Web of Science, and 13 results from Scopus. Upon analyzing and filtering the results and discarding redundancies, we ended up with only seven app evaluation rubrics as is shown in Table 1 below.

3 Findings and discussion

The findings for this study are displayed in the tables below.

<table>
<thead>
<tr>
<th>NO</th>
<th>RUBRIC NAME</th>
<th>AUTHOR(S)</th>
<th>PUBLICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>App evaluation rubric for students with learning disabilities (LD)</td>
<td>(Ok, Kim, Kam, &amp; Bryant, 2016)</td>
<td>Intervention in School and Clinic 2016, 51(4), 244–252</td>
</tr>
<tr>
<td>7</td>
<td>Evaluation rubric for educational apps</td>
<td>(Lee &amp; Cherner, 2015)</td>
<td>Journal of Information Technology Education: Research, 14, 21–53</td>
</tr>
</tbody>
</table>

The search in three academic databases mentioned above elicited only 7 evaluation rubrics as shown in Table 1 above. Table 1 shows the name of the rubrics, the authors of the rubrics, and the journals their research are published in. Table 2 below shows the types of apps the rubrics are meant to evaluate and their evaluative dimensions.

Due to space limitations, Lee and Cherner’s evaluation rubrics (2015) is shown separately below in Table 3 due its extensive categories and dimensions.
**Table 2**: App evaluation rubrics’ targeted apps, evaluation dimensions and rating scales

<table>
<thead>
<tr>
<th>NO</th>
<th>RUBRIC AUTHORS</th>
<th>TARGETED APPS</th>
<th>EVALUATION DIMENSIONS</th>
</tr>
</thead>
</table>
| 1  | (Ok, Kim, Kam, & Bryant, 2016) | Apps catering to students with learning disabilities | **13 dimensions**: objective, strategy, examples, practice, error correction and feedback, error analysis, progress monitoring, motivation, navigation, visual and auditory stimuli, font, customized settings, content error and bias.  
**3 rating scales**: 1 to 3 (3 is best) |
| 2  | (Walker, 2011; Walker, 2013)  
Introduced in 2011, validated in 2013. | All educational apps | **7 dimensions**: curriculum connection, authenticity, feedback, differentiation, user friendliness, student use, student performance.  
**4 rating scales**: 1 to 4 (4 is best) |
| 3  | (Papadakis, Kalogiannakis, & Zaranis, 2017) | Apps for preschool children | **4 sectors, 19 dimensions**:  
**Sector 1: Educational content**  
Dimensions: knowledge package appropriateness, learning provision, levelling, motivation/engagement, error correction/feedback provision, progress monitoring/sharing, bias free.  
**Sector 2: Design**  
Dimensions: graphics, sound, layout/scenery, app/menu design.  
**Sector 3: Functionality**  
Dimensions: child-friendliness, autonomy, instructions existence, configuration ability.  
**Sector 4: Technical characteristics**  
Dimensions: performance and reliability, advertisements/electronic transactions, social interactions.  
**4 rating scales**: 1 to 4 (4 is best) |
| 4  | (Buckler & Peterson, 2012)  
Adapted from Walker’s rubrics | Apps for adults with special needs | **6 dimensions**: application, feedback, adjustability, ease of use, cost, benefits  
**4 rating scales**: 1 to 4 (4 is best) |
| 5  | (Israelson, 2015) | Apps for early literacy education | **4 dimensions**: multimodal features, literacy content, intuitiveness of app navigation, user interactivity.  
**4 rating scales**: 1 to 4 (4 is best) |
| 6 | (Chen, 2016) | Apps for adult learners of English | **7 dimensions:** content quality, pedagogical coherence, feedback and self-correction, motivation, usability, customization, sharing. 

**3 rating scales:** least suitable, average, most suitable |
|---|---|---|---|
| 7 | (Lee & Cherner, 2015) | All educational Apps | **3 categories, 24 dimensions:** The categories and domains are listed in Table 3 below. 

**5 rating scales:** 1 to 5 (5 is best) |

| **Table 3:** Lee & Cherner’s (2015) rubric evaluation categories and dimensions |
|---|---|---|---|
| **Category A: Instruction** | **Category B: Design** | **Category C: Engagement** |
| A1: Rigor | B1: Ability to save progress | C1: Learner control |
| A3: Connections to future learning | B3: Screen design | C3: Pace |
| A4: Value of Errors | B4: Ease of use | C4: Personal preferences |
| A5: Feedback to teacher | B5: Navigation | C5: Interest |
| A6: Level of learning material | B6: Goal orientation | C6: Aesthetics |
| A7: Cooperative learning | B7: Information presentation | C7: Utility |
| A8: Accommodation of individual differences | B8: Media integration | |
| | | B9: Cultural sensitivity |

In Table 3 above, for each evaluative dimension, there is a prompt in the form of a central question which needs to be answered. The rubric for that dimension gives “five indicator descriptors that describe the ways in which an app’s functionality or design may behave in response to a prompt (Lee & Cherner, 2015)”. The evaluators will then rank the descriptor on a 5 point Likert scale where 5 is the highest quality and 1 is the lowest. There is also a ‘not applicable (N/A) option because an app might not be able to be evaluated on all dimensions.

### 4 Conclusions and recommendations

Out of the seven research-informed educational app evaluation rubrics available, it is found that five of them are limited in their use. They either are meant to evaluate apps based on age group such as preschoolers, primary school students, or adult learners of English (Papadakis et al., 2017; Israelson, 2015; Chen, 2016) or based on special needs such as students with learning disabilities or adults with special needs (Ok et al., 2016; Buckler & Peterson, 2012). Only two educational app evaluation rubrics are meant to evaluate all types of educational apps. They are the rubrics by Walker (2011) and by Lee and Cherner (2015).

Analysis of these two rubrics reveal that only Lee and Cherner’s (2015) rubrics are comprehensive enough to effectively evaluate all types of educational apps. Furthermore, the rubrics is also current and up-to-date where it also takes into account today’s approaches to teaching and learning. For example, evaluators can determine whether the
evaluated educational app has features of 21st century learning or collaborative learning embedded in it.

Therefore, to help educators search for and determine quality apps to use in the classroom or to give to their students as out of school activities, we recommend the use of Lee and Cherner’s (2015) educational app evaluation rubrics due to the reasons stated above.

Acknowledgement

This work was financially supported by Universiti Teknologi Malaysia under the Research University Grant (TIER 2) and Ministry of Higher Education Malaysia.

References


