There's an App For That?: Are Mobile Music Games Serious Educational Tools

Charlotte Pierce
The University of Melbourne
Grattan Street,
Parkville 3010,
Victoria, Australia
charlotte.pierce@unimelb.edu.au

Clinton J. Woodward, Anthony Bartel
Swinburne University of Technology,
John Street,
Hawthorn 3122,
Victoria, Australia
cwoodward@swin.edu.au, abartel@swin.edu.au

ABSTRACT
Music students face a significant cognitive load, which often causes them to abandon musical studies. Serious games offer a solution to this problem: present educational content in a fun package to increase student engagement and foster self-regulated, independent learning. In this paper we examine serious music games, specifically on the iOS platform. We address three questions: whether these games exhibit the benefits that serious games are considered to have; whether they provide educational value; and whether they offer any improvement over traditional teaching tools.

We found that although they can offer the benefit of immediate, automated feedback, the currently available games cover only a small amount of musical knowledge. They also tend to support rote-learning style tasks, resulting in low-level learning outcomes, and do not tailor content to players. Despite these drawbacks the games offer some educational value. However, there is significant scope for continued development in the future.

Keywords
serious games, educational games, music education, mobile games

INTRODUCTION
The acquisition of musical skills presents many challenges, both to educators and students (Kwalwasser, 1955). As long as the field of music education has existed there have been debates regarding the best methods of instruction, and although many strong arguments have been made for particular teaching styles, no single method has emerged as the standard. These debates have only been fueled with the increasing availability of technology and personal computing devices (Owens and Sweller, 2008; Konecki, 2015).

A popular use of technology in education is in providing access to ‘edutainment’ (Egenfeldt-Nielsen, 2007) or ‘serious games’ (Sawyer and Rejeski, 2002). These are games which “merge a non-entertaining purpose with a video game structure” (Djaouti
et al., 2011). That is, they are video games which are specifically intended and designed to teach the player knowledge or skills that can be applied outside of the game context.

There are several reasons for the popularity of games as educational tools, both in music and other domains:

- Games are, by definition, designed to engage and motivate players. This means that players are more likely to spend time with a game, and are more likely to enter a flow state when playing. Both of these are significant factors in skill and knowledge acquisition (Elliot and Dweck, 2013; Graesser et al., 2009; Kiili et al., 2012; Ritterfeld et al., 2009).
- Games have the benefit of being perceived as fun, which tends to attract people to play them in the first place. This perception additionally fosters a positive and open mental state in players, which increases their capacity to learn (Greene and Noice, 1988; Pekrun et al., 2002).
- Games are a familiar paradigm to many people, which increases the accessibility of educational content that is presented within the structure of a game.
- Games offer the ability to provide immediate, automated feedback. Given that feedback is a critical component in effective learning (Shute, 2008; Ericsson et al., 1993; Wilson, 2004), this potentially represents a significant benefit of games over traditional education, where students might have to manually look up answers, wait for a teacher to mark their work, or simply guess at their correctness.
- Educational games have been found to encourage self-regulated and independent learning behaviours (Gee, 2003). This means that players take an ownership of their learning, increasing their engagement with the content.
- Games-based learning has been shown to provide valuable context for more formal styles of learning (Cassidy and Paisley, 2013; Vygotsky, 1978; Wood et al., 1976). This context results in more players transitioning to or complementing game-based education with formal learning activities, and progressing faster in their studies.

In this paper we will examine serious games for music education, specifically those on the mobile iOS platform. Mobile games (a.k.a., ‘apps’) were selected as the focus for two reasons. First, tablet applications and devices are becoming increasingly popular as teaching aids due to their mobility, ease of use, and interaction modes. Second, many of the available desktop and web applications have iOS counterparts which have either feature parity or more functionality than the desktop versions. Within the mobile space, the iOS platform was chosen as it features the largest library of serious music games compared to other platforms such as Android.

Our examination of these games will focus on three key questions:

1. Do they exhibit the benefits that serious games are generally considered to have?
2. Do they provide educational value?
3. Do they represent an improvement over traditional tools and resources used in music education?

The next section will describe the method we have used to address these questions. Our findings for each question will then be presented. The final section will provide a discussion of these findings, including potential avenues for future work both in the study and development of serious music games.
METHOD

Game Selection
As this review focused on games available on the iOS platform, the iTunes app store, maintained by Apple Inc., was used to select a representative sample. First, the following search phrases were defined:

- music theory
- music theory tutor
- music tutor
- sight reading
- music sight reading

A recursive process was then applied to expand the results of searching these phrases in the app store into a larger list of apps to consider for inclusion in the study. Given the limited number of results the iTunes app store returns for individual search queries, this process heavily relies on Apple’s algorithm for identifying related apps.

After searching a phrase, each game in the results list was considered in turn. If the game was not already in the list of candidates, it was searched. Searching involved adding the game to the list of candidates, opening the list of games related to it, then searching each of the games from that list that had not already been seen. This process iteratively expanded the space of connected apps, meaning a large number of candidates could be identified from only a small number of search phrases.

Once a list of candidates was found, it was filtered according to the criteria illustrated in Figure 1. The first key selection criterion was that the application must be available on the Australian iOS app store, as this is the only version of the store accessible from Australia. If it passed this test, the game’s description and screenshots were perused in an attempt to identify whether it is a skill-based game (defined by Cherner et al. (2014) as being a game which aims to help a learner build basic abilities and fundamental knowledge in a subject area). Once past this stage, the game’s rating and reviews (both positive and negative) were briefly considered. This was done in an effort to identify any obvious issues (e.g., frequent crashing, catastrophic bugs) that would prevent the game from being properly assessed. In cases where a game clearly had such issues it was omitted. Finally, the game was checked for compatibility with the review hardware.

This process of selection and filtering resulted in a final sample of 175 out of 313 games being selected.

Measuring Content Coverage
In order to describe the content of the games in a consistent, neutral language, a taxonomy of musical content and activities was used. Proposed by Pierce and Woodward (2017), the taxonomy comprises content areas and activities. Content areas are defined as the categories into which musical knowledge can be placed. Each category contains a subset of topics, all of which relate to the same broad area. Activities are ways in which the knowledge in content areas can be applied. Figure 2 shows the structure of the taxonomy as a hierarchical graph.

This taxonomy was used as a framework for describing the games. Each part of every game was matched to either a content area or activity from the taxonomy. This was
Figure 1: The process used to filter candidate iOS games for inclusion in the study, using four key criteria

Figure 2: A hierarchical overview of the taxonomy used to describe the games. Taken from Pierce and Woodward (2017).
done with a high level of granularity, to the point where games were described on a level-by-level basis.

Once the parts of a game were matched to appropriate content areas and activities, the entire game was rated for its depth of coverage and level of focus in each part of the taxonomy. In accordance with recommendations made by Lee and Chernher (2015) in their work developing rubrics for assessing instructional applications, ratings followed a five-point Likert scale, using the categories ‘Very low’, ‘Low’, ‘Medium’, ‘High’, and ‘Very high’.

**Identifying Learning Outcomes**

In order to understand the depth of learning players are likely to achieve, the learning outcome of each game was identified. The learning outcomes considered were those from the SOLO (Structure of the Observed Learning Outcome) taxonomy developed by Biggs and Collis (1982). According to this taxonomy, five learning outcomes are possible:

1. **Pre-structural**: Indicates no understanding of the topic; the student hasn’t understood the point of the task and uses simple methods for completing it.
2. **Uni-structural**: Indicates understanding of one aspect relevant to the topic; the learner has a basic understanding.
3. **Multi-structural**: Indicates understanding of multiple aspects relevant to the topic, but in isolation; assessment is mostly quantitative.
4. **Relational**: Indicates an ability to connect multiple aspects relevant to the topic; student somewhat understands the topic.
5. **Extended abstract**: Indicates an ability to generalise or extend knowledge of the topic to a new area.

**Identifying Feedback Mechanisms**

Feedback, defined by Mason and Bruning (2001) as being any message provided in response to a learner with the purpose of helping them identify their errors and misconceptions, is a key factor in effective learning (Shute, 2008; Ericsson et al., 1993; Wilson, 2004). Two types of feedback styles were originally defined by Kulhavy and Stock (1989):

1. **Verification**: Indicating to the learner whether their answer is correct or incorrect.
2. **Elaboration**: Providing the learner with cues to lead them towards a correct answer.

Elaboration was later split into two variations by Shute (2008):

1. **Directive Elaboration**: Addresses the learner’s response in terms of their particular errors. Also known as specific elaborative feedback.
2. **Facilitative Elaboration**: Provides worked examples or information related to the topic. Also known as general elaborative feedback.

The feedback style used by each game considered in this study was identified as being verification or elaboration and, if elaboration, which variation.

The feedback type used by each game was also identified. This is a more descriptive label than the feedback style, as it formalises how the correctness of an answer is communicated and what (if any) additional information is provided to the student. The possible feedback types are the twelve defined by Mason and Bruning (2001) and Shute (2008):
1. **None:** No feedback is given.
2. **No-feedback:** Individual answers are not addressed. For example, the learner is given an overall score but is not told which questions they got correct or incorrect.
3. **Knowledge-of-response:** Individual answers are labelled as correct or incorrect.
4. **Answer-until-correct:** Individual answers are labelled as correct or incorrect, and if incorrect the user is required to try again.
5. **Knowledge-of-correct-response:** Individual answers are labelled as correct or incorrect, and if incorrect the correct answer is provided.
6. **Error-flagging:** Mistakes in individual answers are highlighted, but not explained or corrected.
7. **Topic-contingent:** Individual answers are labelled as correct or incorrect, and if incorrect the user is directed towards materials that will help them identify the correct answer.
8. **Response-contingent:** Individual answers are labelled as correct or incorrect, and if incorrect the correct answer is provided. Explanations are given in all cases – for example, “Your answer was incorrect because ...”; “Your answer was correct because ...”.
9. **Hints-cues-prompts:** Individual answers are labelled as correct or incorrect, and if incorrect hints are given as to the correct answer. The correct answer is not explicitly provided.
10. **Bug-related:** A ‘bug library’ is maintained, and examples are retrieved relating to a learner’s error(s) when they give an incorrect answer.
11. **Attribute-isolation:** Individual answers are labelled as correct or incorrect, and the central attributes of the target concept are highlighted.
12. **Informative-tutoring:** Individual answers are labelled as correct or incorrect, and, if incorrect, mistakes are highlighted and hints as to the correct answer are given. The correct answer is not explicitly provided.

**Identifying Progression Models**

The progression model defines how users work through the content and activities of a game. Only one progression model can be used at any one time. Four progression models are defined:

1. **None:** There is no content to progress through. For example, a game with one level.
2. **Linear:** There is only one path through the content.
3. **Manually adaptive:** The player can alter some or all aspects of the content. For example, a quiz game which allows the player to select the topics and concepts they will be quizzed on.
4. **Adaptive:** The game automatically alters some or all aspects of the content. For example, a quiz game which presents more questions from subject areas in which the user has performed poorly.

**RESULTS**

**Mobile Music Teaching Games as Games**

From a broad perception level, it is interesting to note how these games are categorised in the app store. Categorisations are self-selected by developers, so indicate how the developers both perceive their work and wish it to be perceived by potential users. The genre of an application determines where it is contained the store, the charts it appears in, and its prominence in search results (Apple Inc., 2019). Apple recommends that developers consider an application’s purpose, where users may look for that type of
application, and where similar applications are placed in the store before selecting genres for their work.

Every application in the store has a primary genre. This can be further specified with secondary, tertiary, and quaternary genres. Not all applications have four levels of categorisation, but all have nominated at least primary and secondary genres.

In the games that were sampled, a core set of genres were consistently repeated at all levels of categorisation: ‘Music’, ‘Games’, ‘Education’, and ‘Educational’. Other genres such as ‘Reference’, ‘Entertainment’, and ‘Family’ also appear, but less frequently. Most genre sequences contain only a small number of games. However, both the ‘Music → Education’ and ‘Education → Music’ sequences describe a large portion of the sample, containing 63 and 48 games respectively. This supports the idea that developers see their work as contributing to the world of educational (i.e., serious) games for music.

Although the developers of these games may see them as serious games, this does not necessarily mean that they provide the benefits that serious games are generally considered to offer. One of the supposed benefits they do clearly exhibit is that of immediate, automated feedback. Nearly all of the games provide the user with feedback on their performance. Furthermore, the feedback is usually immediate (i.e., the user is told whether their answer is correct as soon as they submit/select it). The few games which don’t provide feedback either do not contain activities (i.e., there is nothing to give feedback on), or were deliberately designed to encourage self-reflective feedback.

Whether the games are engaging and motivating can be less clearly stated. They do, however, generally provide scoring systems which inherently challenge players to improve on their past performances (Toups et al., 2009; Cagiltay et al., 2015). Some provide more explicit incentives to improve, such as online leaderboards.

Many of the games also offer the ability for players to self-pace or self-manage their learning. This is often in the form of selecting difficulty levels or curating the content of the activities.

Overall, the games do offer at least some of the benefits attributed to serious games. Other potential benefits, such as whether they encourage transitions to formal musical studies, or whether they are perceived as being fun, warrant further investigation.

**Mobile Music Teaching Games as Educational Tools**

As described in the method, each part of every game was matched to either a content area or activity, using a taxonomy for musical content and activities proposed by Pierce and Woodward (2017). Using this taxonomy, it was found that **Reading** and **Instrument-specific** are the most popular content areas seen in the sample of games. These are defined as follows:

**Reading Content:** covers the aspects of musical knowledge needed to understand, from musical notation, what pitch to play and its intended length relative to other pitches. These are the most basic elements needed to visually communicate a sequence of notes – both monophonic and polyphonic – in such a way that they can consistently be reproduced. For example, note types, rest types, the stave, clefs, and accidentals.

**Instrument-specific Content:** covers information specific to the playing of a particular instrument. This is one of the broadest content areas. For example, hand positions, tuning, and embouchure.
Table 1: Coverage of content areas across the sample of games

<table>
<thead>
<tr>
<th>Category</th>
<th>Covered?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Reading</td>
<td>142</td>
</tr>
<tr>
<td>Rhythm</td>
<td>158</td>
</tr>
<tr>
<td>Scales</td>
<td>159</td>
</tr>
<tr>
<td>Elements of Harmony</td>
<td>155</td>
</tr>
<tr>
<td>Harmonic Structures</td>
<td>169</td>
</tr>
<tr>
<td>Style</td>
<td>168</td>
</tr>
<tr>
<td>Instrument-Specific</td>
<td>140</td>
</tr>
<tr>
<td>Historical and General Knowledge</td>
<td>173</td>
</tr>
</tbody>
</table>

That these are the most popular content areas means that most of the games focus on teaching either basic note reading or specific instrument-related skills. The number of games covering each content area is shown in Table 1.

The most popular activity type was Visual recognition – Table 2 shows that 87 of the games contain activities of this type. These are activities where players are asked to identify some musical element, and question content is largely visual. For example, players might be shown a picture of four different note types and asked to identify which one is a quaver.

Table 2: Coverage of activity types across the sample of games

<table>
<thead>
<tr>
<th>Category</th>
<th>Covered?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Auditory Recognition</td>
<td>133</td>
</tr>
<tr>
<td>Visual Recognition</td>
<td>87</td>
</tr>
<tr>
<td>Auditory Description</td>
<td>166</td>
</tr>
<tr>
<td>Visual Description</td>
<td>161</td>
</tr>
<tr>
<td>Visual Playback</td>
<td>121</td>
</tr>
<tr>
<td>Memory Playback</td>
<td>163</td>
</tr>
<tr>
<td>Notation</td>
<td>149</td>
</tr>
</tbody>
</table>
The games were also given an overall rating for their depth of coverage and level of focus in each part of the taxonomy. From this we found that the games tended to have a high focus, but a low depth. This means that they typically touch on just one or two concepts, and cover only the basics of those concepts.

It is mostly because of this low depth, and the heavy use of visual recognition activities, that 132 of the 175 games have only uni-structural learning outcomes. As defined in the method, this is a learning outcome where learners obtain a basic understanding of a single aspect of knowledge. This is a typical learning outcome when students are only given rote-learning styled activities involving simple identification tasks, as many of the games do. Visual recognition activities are most likely popular as they are easy to implement in software, and it is trivial to assess player’s answers algorithmically. A small number of the games operate at a multi-structural level, and the remainder of the sample operate at a pre-structural level. The relational and extended abstract learning outcomes are not seen in any of the games.

These low-level learning outcomes are supported by the feedback mechanisms seen in the games. It is most common for the games to provide feedback immediately. By combining this information with that in Figure 3, which shows that the large majority of games utilize the Verification feedback style, it can be said that most of the games immediately indicate to players whether their answer is correct or incorrect, but provide no additional guidance. This is further supported by Figure 4. This shows that the most common feedback types seen in the games are Knowledge-of-response, Knowledge-of-correct-response, and Answer-until-correct. These types are all similar, in that they all indicate to players whether they are correct or incorrect. They differ only in whether they stop at that point (i.e., with Knowledge-of-response), provide the player with the correct answer (i.e., with Knowledge-of-correct-response), or require the player to keep trying until they discover the correct answer for themselves (i.e., with Answer-until-correct). None of these feedback types offer additional guidance such as materials to help players understand why their answer was incorrect, or to assist them in finding the correct answer.

Put together, these results show that whilst the games typically provide players with feedback on their answers, this feedback is simple and rarely goes beyond simply telling them whether they were correct or incorrect. Feedback does not extend to showing players how they are performing over time, or how their performance compares to other players. Players are rarely given any indication on how they could improve their skills, or details on where and how their answer was wrong. At most they are invited to continue answering until they are correct.
Although limited feedback can be a mechanism for encouraging students to find answers on their own (Mason and Bruning, 2001), in this case it is more likely to be evidence that developers are selecting the simplest option. More sophisticated methods of feedback would require gaining a perspective of how and where a player went wrong, what gap in their mental model this reveals, and how to effectively address that gap. This is difficult and time-consuming to implement.

In terms of progression through content, most of the games offer players a manually-adaptive style. Figure 5 shows the distribution of progression styles used, revealing that games which do not use a manually-adaptive style tend to have no progression or follow a linear model. Exactly one game in the sample – Music Reading Essentials (Apricot Digital Publishing, 2018) – is described as having an adaptive progression style. Music Reading Essentials contains a finite set of pre-written questions and uses the number of times a player has answered a question incorrectly to determine the frequency with which that question should be repeated. This approach relies on the existence of a set of pre-defined questions. As such, this implementation of the adaptive progression style does not represent a particularly significant enhancement over the manually adaptive alternatives.

The decision to use a manually-adaptive progression model is a double-edged sword. On the one hand it does allow players to tailor the content of the game, theoretically so that they can target areas where they are weak. However, this is a task that players may not be suited to simply because they are still learning the content. That is, they may not have the knowledge and experience to set appropriate parameters.

As with the most popular activity type, feedback type, and feedback styles, the almost unilateral use of a manually-adaptive progression model might again indicate a choice by developers to simply do what is easiest to implement. By making users select the content they will be shown, developers are offloading the mental effort of deciding when it is appropriate to present learners with certain concepts and what order they should be presented in. All the developer needs to provide instead is a basic user interface for setting parameters.

Overall, the games can be said to act as educational tools, but at the same time their utility is limited. They offer mostly rote-learning activities, and focus on teaching musical notation reading and basic instrument-specific skills. This, and the feedback mechanisms used, result in generally low-level, shallow learning outcomes.

Mobile Music Teaching Games as Compared to Traditional Teaching Tools

Compared to traditional teaching tools in music education, serious games can offer some attractive benefits. Aside from being potentially more engaging, motivating, accessible, and generally fun than traditional approaches, serious games can provide players with immediate feedback on their performance. This is the main way in which the currently available games appear to represent an appreciable improvement.

However, the current selection of games are also limited in their utility. As discussed in the results, at this time they offer players low-level learning outcomes, and cover

---

---

---
only a limited selection of musical knowledge. According to Hein (2014), one of the primary difficulties in learning music is that students need to gain mastery in three areas:

1. Musical concepts,
2. Musical notation, and
3. Instrumental techniques (i.e., translating notation into sound).

Of these areas, most games target the learning of musical notation, some games help with learning musical concepts, and few games instruct players in instrumental techniques. There are also additional areas not included by Hein which are covered by few to no games, such as improvisation and musical interpretation. Because of this, at best serious music games should probably only be used to augment traditional teaching, or to provoke an interest in broader musical learning. They can not on their own provide an education in fundamental musical knowledge and skills.

**DISCUSSION AND FUTURE WORK**

Music students, particularly beginners, face a significant cognitive load – one that often causes them to disengage or simply abandon their studies (Mota, 1999). Serious games offer a solution to this problem: present content in a fun, attractive, familiar package to increase student’s engagement and help them foster self-regulated and independent learning behaviours. However, although these are the advertised benefits of serious games, whether the available games are actually providing these benefits is another question.

Overall, the currently available serious games for music on the iOS platform tend to be narrow in focus, shallow in depth, and simple in design. In every aspect, developers seem to have trended towards making choices that simplify and reduce development time and complexity. This has led to a lack of depth in covering musical knowledge, low-level learning outcomes, and lack of sophisticated feedback across the sample. It also means that there are no games which tailor their feedback or content to the player. Instead, they rely on allowing players to manually set the parameters of the content they will see.

Together, these results combine to reveal a large number of gaps in the field. For example, many of the games could make use of more sophisticated feedback types. For example, if a player is asked to identify the B major key signature from a list of four options and selects the wrong one, instead of simply telling the player they are wrong the game could show them how to determine the number of sharps or flats in any key signature.

Progression models across the sample are limited, with the most sophisticated offerings being those where players can set the parameters of the content. Unfortunately, this is a task players may not be suited to given that they are still learning the content. Adaptive progression models present a significant improvement over this, as content would be automatically tailored to the player’s individual misunderstandings. Many of these gaps could be solved algorithmically to great effect, but this has not yet been done.

In terms of the areas of knowledge covered, whilst notation reading skills are heavily represented many other areas are ignored. For example, identifying chord progressions within existing music is rarely covered, nor is creating aesthetically pleasing progressions from scratch. No game covers the principles of melody writing, or instrumentation and orchestration. Instrument-specific instruction rarely goes beyond
identifying single notes on an instrument or providing fingering charts for specific notes or chords.

These areas all represent rich avenues for future work, which would result in higher quality games and players who achieve higher-level learning outcomes. It would also be making more effective use of the capabilities of modern technology, which are not currently being exploited. Over time, serious games for music should only improve in quality, content coverage, and educational value.

BIBLIOGRAPHY


Wilson, T. C. (2004). An implementation of a drill and practice system to assist in the teaching of basic music theory.


ENDNOTES