Motivating Students for Ear-Training with a Rhythmic Dictation Application

Matevž Pesek 1,*, Lovro Suhadolnik 1, Peter Šavli 2 and Matija Marolt 1

1 Faculty of Computer and Information Science, University of Ljubljana, 1000 Ljubljana, Slovenia; ls2964@student.uni-lj.si (L.S.); matija.marolt@fri.uni-lj.si (M.M.)
2 Conservatory of Music and Ballet, Ljubljana, 1000 Ljubljana, Slovenia; peter.savli@kgbl.si
* Correspondence: matevz.pesek@fri.uni-lj.si

Received: 27 July 2020; Accepted: 24 September 2020; Published: 27 September 2020

Abstract: We present the development and evaluation of a gamified rhythmic dictation application for music theory learning. The application was developed as a web application in the existing Troubadour platform for music ear training. The focus of the developed application was on user experience and engagement. With regards to the former, we developed a responsive and intuitive user interface; for the latter, we included several gamification elements and assessed the impact on the students’ engagement. We report on students’ experience analysis done through questionnaires and background data collected through the platform. We evaluated the rhythmic dictation application with the conservatory-level music theory students through A/B testing in order to independently evaluate the application’s impact. The results show a significant impact of the application on the students’ exam scores.

Keywords: music theory education; ear training; rhythmic dictation; music information retrieval; gamification; e-learning; user evaluation

1. Introduction

A number of different problems have been researched in signal processing and music information retrieval (MIR) communities. Some topics, such as music transcription (e.g., [1–3]), artist similarity, and recognition (e.g., [4,5]), have been dealing with information retrieval and classification, while others—including the automatic accompaniment (e.g., [6]), music theory learning, and automatic generation and evaluation of music exercises (e.g., [7,8])—have been focused on aiding the learning process of performers, writers, and music creators. In general, the music theory learning process requires talent, patience, and a significant amount of time to practice. Consequently, music theory learning and ear training are unpopular, especially among children and students. To aid the learning process, gamification of learning has taken a significant step towards easing the traditional methods of learning [9,10]. With gamification, the user perceives the task at hand as less serious and its completion more gratifying. The goal of a gamified system is to give the user well-balanced tasks and reward them by positive reinforcement elements in order to gain and retain user’s engagement. It could also be argued that modern gamification techniques are rather digitized than invented [11]. Coming from the evolving field of computer game production, gamification was re-discovered through “serious” games and game-based learning as digital techniques for user engagement [12]. Consequently, much work has been invested in researching the underlying processes for students’ engagement [13] and several products for e-learning have been introduced—from learning management systems, such as Moodle (https://moodle.org), to specialised platforms and apps [14].

In this paper, we present a rhythmic dictation application in an open-source web platform. The application incorporates the automatic generation of exercises on a variety of difficulty levels, with
adjustable difficulty according to the curriculum within those levels and several gamification elements
to gain students’ engagement. The application and the Troubadour platform are developed as a web
service to avoid the need for (native) mobile development and ease future development with the most
commonly used web technology currently available. The platform is also free for non-commercial
use and, therefore, suitable for in-class use by public schools and individuals, without the additional
financial burden. It offers a supportive and flexible environment for the teachers and the students.
The goal of this paper is to assess the developed application primarily in terms of students’ performance
and its impact on the conventional exam results.

2. Related Work

During the last decade, a variety of gamified e-learning approaches have been proposed. In their
overview, Antonci et al. [15] analyzed the gamification approaches across variety of fields and discussed
several common gamification elements used across different platforms, such as badges, leaderboards, time
limits, points, challenges and feedback. Similarly, Gallego-Duran et al. [16] developed a ten-characteristic
game-design-based gamification rubric, which also included—besides the aforementioned challenge and
feedback—randomness, automation, open decision space, learning by trial and error, progress assessment,
emotional entailment, and playfulness. Grant and Betts [17] analyzed the influence of badges to motivate
the production of good content through submission, editing, and review. Similarly, Hakulinen et al. report
the students while using a badge system in an learning environment gave positive feedback, had a higher
number of sessions, and spent more time in total in the learning environment, and were, therefore, engaged
by the badge system. De-Marcos et al. [18] tested social networking and gamification with elements of
badges and leaderboards to motivate achievement, collection and competition. Their results report better
performance, as compared to the traditional e-learning approach for practical assignments, but not for
assessing knowledge. Huang and Hew [19] evaluated whether points, badges, and leaderboards affect
learning and activity. Their results also showed that gamification was helpful in cognitively engaging
learners and increasing their activity. However, the students’ knowledge assessment through traditional
exams was not included in the study.

All of the above mentioned gamification aspects and rubrics are somewhat universal and
many of them can also be directly applied to music learning and ear training. In recent years,
several approaches have been proposed for music learning and training. During the last decade,
a variety of instrument learning applications have taken the form of mobile device applications
available on both Google Play and Apple App stores for both mobile platforms. Several applications
also incorporate gamification elements. Among the different instruments, piano learning is perhaps
the most well-represented. One of the most popular gamified platforms for instrument learning is
Yousician (https://yousician.com/—also available on App and Play stores) [20]. Similarly, popular is
the Simply piano application (https://www.joytunes.com/—also available on App and Play stores).
Synthesia (https://synthesiagame.com/) [21] visually represents the piano roll and, therefore, eases
the first steps into piano practice without the need for the user’s knowledge about music notation.
A project called ForteRight, among other similar approaches, also proposes a hardware piano strip for
a similar purpose. Another popular option for piano learning is My Piano Assistant (Available on Play
and App stores).

There is a variety of applications supporting the guitar and wind instruments training. Besides Yousician, which also supports guitar learning, a highly gamified application for guitar learning
Rocksmith (https://www.ubisoft.com/en-us/game/rocksmit) offers an experience that is similar to
the game Guitar Hero and is also available for PC and console users (Microsoft XBOX 360 and Sony Play
Station 3). Tonestro (https://www.tonestro.com/) (wind instrument learning application) offers similar
experience for flute learners. The Finger Charts application offers flute, saxophone, oboe, and clarinet
players a tool to practice finger positions. There are several applications that are currently available
for music theory and ear-training. The Tenuto (https://www.musictheory.net/products/tenuto—also
available on App store) application offers 24 types of quiz-like exercises for music theory. Similarly,
the ABRSM Theory Works application provides a quiz-like environment for music theory practice. Some applications tend to include user profiles and achievement sharing in their platforms. For example, MusicTheoryPro (http://musictheorypro.net)—also available on App store) application incorporates elements of user profiles and achievement sharing among the application users.

However, most of the available commercial applications are focused on adolescent or adult users. Serdaroglu [22] performed a review of about 60 applications on the Apple App store, available in 2018, researching their potential use for pre-elementary and elementary education. They discovered that only a small subset of the available ear training applications found in Apple app store can be used for younger children.

Gamified music-related applications have also emerged as a medium for collaborative data gathering [23]. Specifically in music information retrieval, several approaches for gamification of music annotation and meta-data gathering have been used. Uyar et al. [24] proposed a rhythm training tool, focused on the specific music genre of usuls in Turkish makam music. Kim et al. [25] proposed the Moodswings game for mood labelling, where the users were asked to plot the mood on the valence-arousal graph. They collected over 50,000 valence-arousal point-labels on more than 1000 songs. The authors identified gamification as the key component of user engagement. In a similar manner, Mandel and Ellis [26] proposed a web-based game for collecting song meta-data, such as genre and instrumentation. Law et al. [27] created the TagATune game for music and sound annotation. The game collects comparative information about sounds and music, where users play the game in pairs. The authors collected responses from 54 test users. They also focused on user engagement through three aspects: sense of the user’s competence, pleasantness and sensory user experience, and the opportunity to connect with a partner. Burgoyne et al. [26] presented the game, named Hooked, to explore the “catchiness” of songs on the responses provided by 26 users. The dataset consisted of 32 songs. Aljanaki et al. [28] developed a ‘game with a purpose’ to gather emotion responses to music. They collected more than 15,000 responses from 1,595 participants. Overall, in the MIR community, the developed applications mainly served as a medium to gather data. When considering the related work on this topic in MIR, the proposed games, which were made for gathering data from smaller user groups, commonly did not emphasize the incorporated gamification elements, while the authors of those games, which were developed for a larger scale user base, also reported gamification and user engagement as important aspects for their applications’ success. There were also several non-commercial applications that were proposed for music learning, incorporating gamification to gain engagement and improve students’ performance. For example, Gomes et al. [29] proposed a flappy crab, a clone of the famous flappy bird game, for music learning. Rizqyawan et al. [30] proposed an adventure game for music theory and ear training. They reported a significant increase in performance in theory exams and relative pitch tests on 40 elementary school students.

Currently, the industry’s main focus is (mobile) applications for instrument learning, while there is a smaller number of applications focused on music theory and ear training. In general, the instrument-oriented applications tend to include more gamification elements. Most of the aforementioned approaches are also either commercial or closed-source. Moreover, most of the available commercial applications and services are not free to download and they are most often immutable by an individual to adjust the phases of learning to their curriculum. While these applications can help the user to improve their knowledge and performance, it is more difficult to use them in-class, due to the inability to modify the content without code-level access. Additionally, the use of the non-free applications in-class would financially burden the music schools and individuals.

Summary

There is an opportunity for open-source applications for ear training. With a combination of the MIR available techniques to automatize exercise generation and in combination with gamification, such an application could engage the music theory students in their learning process and positively influence their skills and performance. Several methodological overviews that are referenced in this
paper describe similar gamification aspects, which can be applied to the proposed rhythmic dictation application:

- **Feedback**—the application should automatically evaluate the user’s response and give feedback to the user about their performance instantly.
- **Points**—based on the correctness of the user’s response, the application should calculate achieved points.
- **Badges**—through collecting points and performance (i.e., time played, decrease time needed for completion, etc.), the user should achieve various badges, expressing specific achievements.
- **Levels**—the user should progress through levels, which gradually increase the difficulty of the exercises.
- **Time limit**—the exercises should be time-limited.
- **Progress bars**—the user’s current progress should be displayed to the user to show them their progress and improvement.
- **Avatar**—the user can modify their avatar (image, text) and display collected progress (levels, badges, progress bars) on their profile page.

The work that is presented in this paper introduces an application for rhythmic dictation, which is integrated into the existing e-learning platform for aural training, Troubadour [31]. It is implemented as a web-based application with a responsive user interface, which is specifically designed for mobile devices, the most commonly used electronic device among the students. The application automatically generates exercises according to the individual student’s level of knowledge and in-app progress. Gamification elements, including badges and leaderboards, are implemented, in order to increase student’s engagement. The source code for the platform and the rhythmic dictation application is publicly available, future development, and research. Furthermore, we describe the development of the rhythmic dictation application. Subsequently, we analyze four aspects of the application’s development and in-class use with the first and second year conservatory-level students: the newly-developed interface for rhythmic elements input, the automatic exercise evaluation, the students’ engagement, and the application’s impact on students’ performance.

3. The Rhythmic Dictation Application

We developed a rhythmic dictation application, which was incorporated into the Troubadour platform (https://trubadur.si). The Troubadour platform is a framework for music theory learning with the support for gamification elements, including badges, points, and leaderboards. The application and the platform are easily deployable with the use of package management tools, and the code is available as open source software and publicly accessible on Bitbucket (https://bitbucket.org/ul-fri-lgm/troubadour_production).

When considering the most common types of exercises performed by the music theory students—melodic (interval) dictation exercises, rhythmic dictation exercises, and harmony exercises—the rhythmic dictation application covers one of the three fundamentals of Western music theory curriculum. The conventional practice, performed in-class and at home by the students nowadays, includes listening to a pre-recorded or teacher-performed dictation and solving it on paper. The exercise evaluation and grading done by teachers usually takes a few days. Homework means additional work by teachers, who have to upload recordings, digitized in a separate tool (such as Sibelius), on a web-site or a school LMS web-portal. Evaluation of the exercises also takes a significant portion of the teacher’s time, depending on the amount of provided exercises and number of students. From the student’s point of view, the feedback on their work is given considerably late, probably potentially influencing the student’s learning process. Instant feedback could be beneficial for the student, but impossible to achieve by the teacher, unless the class is performed individually. The rhythmic dictation application offers an easy to use and automated way for students to solve rhythmic dictation exercises in-class and out-of-class with the immediate feedback on their performance and the customizable exercise generator that adapts the
difficulty of generated exercises to the student’s level of knowledge. In addition, it lowers the teachers’ overhead work with the conventional exercise generation and evaluation.

3.1. Technical Details

The application was developed as a responsive web application, which adapts well to mobile devices. In this way, the development and maintenance of the platform is simplified, as the platform is browser-accessible on all major platforms—Windows, Linux, OS X for desktop environments, as well as Android and iOS for mobile devices. The platform offers intuitive upgradeability of applications within the platform by expanding the educational platform with new types of applications. The platform applications include several gamification features that are shared across the applications. Finally, the developers focused on the enhanced user experience by including customization of the exercises based on the user’s personal preferences, as well as customizable exercise difficulty according to educational level.

The platform is implemented in PHP with Laravel framework. The front-end, which includes the views to included applications, consists of two modules: the administrative module, which includes registration and authentication and uses the Laravel Framework Blade templates; while the second module, which consists of the exercise and gamification automation logic, is implemented with Vue.js. All communication with the web server after authentication and initialization of the basic component of the Vue.js framework is done via asynchronous web requests to the API server, which returns data in the JSON format. This methodology aids future development of the platform and enables changes of individual modules on the front-end and back-end, if needed in the future.

3.2. The User Interface

The user interface of the rhythmic application follows the conventional steps in practice. The students listen to a rhythmic sequence, which they have to write down in music notation. Therefore, the main part of the rhythmic dictation application’s user interface (Figure 1a) includes two staves displaying the input rhythmic sequence and a rhythm input interface. The upper (smaller) stave shows the entire sequence with a red rectangle indicating the area shown in the lower larger stave, where the user inputs their response to the dictation. The dictation can be repeatedly played-back and paused while playing.

Each exercise begins with a metronome indicating the meter and is followed by the rhythmic dictation playback. The student can pause and replay the dictation, and adjust the playback speed and volume. The dictation is played while using an organ sound. The sound was chosen due to its fast onset, steady sustain, and a clear offset, in agreement with music theory teachers. While the sound of piano is commonly used for a melodic dictation, its unclear offset can cause ambiguities in determining the event length (vs. pause).

While there are several well-established guide lines for developing a user interface, the specificity of the rhythmic exercises unveiled an intriguing UX challenge. The rhythm input keyboard supports a variety of rhythmic inputs: note and pause lengths, subdivisions, and syncopation. The inputs are split into two layouts to accommodate for the small screens of mobile devices: on the primary layout, the most common note and pause lengths are displayed. With keys for subdivision and syncopation, the layout changes to show a set of additional input options, as shown in Figure 1b. When considering both screen orientations, the input re-adjusts to the landscape orientation for those users, who prefer this device position, as shown in Figure 2. The upper (smaller) stave is removed and the lower stave is enlarged over the upper half of the screen.
Figure 1. The main screen of the rhythmic dictation application on a mobile device (left). The primary rhythm input interface is shown below the staves. The secondary keyboard layout with options for adding and modifying subdivisions is shown on the (right).

Figure 2. The landscape screen orientation layout of the rhythmic dictation application on a mobile device. The input keyboard is re-arranged to aid user experience in this position.

3.3. Automatic Generation of Exercises

The application includes an exercise generator that can generate exercises of different difficulty levels. The difficulty of a rhythmic exercise is governed by several parameters: subdivision complexity (from quarter notes, to 32nd notes), subdivision types (dual vs. ternary), subdivision distributions, and the number of events (length of the sequence). In order to generate meaningful random sequences, the distributions of these parameters must be governed. Otherwise, the generation may result in musically meaningless and unrealistic sequences. The difficulty of such sequences would also vary highly, impacting the students’ engagement. For the initial parameter distribution values, we analyzed the existing Slovenian curriculum materials made and provided by the teachers. With their help, we divided the materials into 16 different difficulty levels, based on the current curriculum. This approach resulted in randomly generated sequences, which are appropriately difficult and
musically meaningful, therefore engaging the students individually with sufficient difficulty, while not overwhelming them with either too difficult, simple, or meaningless examples.

The 16 difficulty levels include difficulty standards in the range of elementary music school, conservatory and academia. The levels are split into four major levels, and each major level is split into additional four minor levels. The levels are marked with numbers [11–14, 21–24, 31–34, and 41–44], where the first digit corresponds to the major difficulty level, while the second digit corresponds to the minor difficulty level. The parameter distributions for each level were set as the default values for exercise generation in the rhythmic dictation application. However, individual teachers can additionally modify the distributions according to their students specific needs.

The in-depth procedure of sequence generation goes, as follows: the first stage of the algorithm contains the initial pre-preparation steps before the generation. First, the rhythmic level \( l \) is chosen from the available levels:

\[
\mathcal{L} = \{11, 12, 13, 14, 21, 22, 23, 23, 31, 32, 33, 34, 41, 42, 43, 44\}.
\] (1)

The algorithm then randomly selects the meter type \( B \) (time signature), on which it generates the sequence. The meter type is defined with \( b_d \), which represents the number of beats, and \( b_b \), which represents the basic beat value. Together, they form any time signature, representing the upper and lower numerals in the signature. Each meter type is also limited to the the lowest rhythmic level \( l_{\text{min}} \), where it can occur.

\[
B = (b_d, b_b, b_{l_{\text{min}}}) \mid b_d \in \mathbb{N} \land b_b \in 2^u \land u \in \mathbb{N} \land b_{l_{\text{min}}} \in \mathcal{L}
\] (2)

The algorithm also includes support for mixed meter types \( B_s \), which are a generalization of basic meter types, where:

\[
B_s = (b_d, b_b, b_{l_{\text{min}}}, K) \quad K \in (k_1, k_2, \ldots, k_i, \ldots, k_s)
\] (3)

and individual meter type \( k_i \) is defined as:

\[
k_i \in \{ (k_{d_i}, k_{b_i}) \mid k_{d_i} \in \mathbb{N} \land k_{b_i} \in 2^u \land u \in \mathbb{N} \}
\] (4)

where the time-length of all \( s \) events in individual meter types \( k_i \) in mixed meter is equal to the time-length of two bars in the first chosen meter:

\[
\sum_{i=1}^{s} k_{d_i} = \frac{b_d}{b_b}.
\] (5)

Each generated rhythmic pattern is represented as a sequence of ordered pairs:

\[
v = (v_{\text{len}}, v_{\text{cross}}), \quad v_{\text{len}}, v_{\text{cross}} \in \mathbb{Q},
\] (6)

where \( v_{\text{len}} \) represents the length of the entire rhythmic pattern in quarters and \( v_{\text{cross}} \) represents the pattern’s length (in quarters) in the next bar. Most rhythmic patterns do not exceed one bar length and have a value of \( v_{\text{cross}} = 0 \). We denote patterns, where \( v_{\text{cross}} > 0 \) as “binding patterns”. These patterns serve as a link between individual segments within the generated sequence. The algorithm selects the binding pattern before generating individual segments in both bars, in order to allocate adequate space. Figure 3 shows a simple binding pattern used by the algorithm. This pattern is defined as \( v = (2, 1) \), where its total length equals to two quarter notes, and its cross length of one quarter note.

A rhythmic exercise with \( t \) beats can then be defined as an ordered sequence of rhythmic patterns

\[
V = (v_1, v_2, \ldots, v_t, \ldots, v_n),
\]

with length \( L \in \mathbb{Q} \) defined by the chosen meter type \( B_s \):
\[ v_i = (v_{\text{len}}^i, v_{\text{cross}}^i), \quad B_s = (b_d, b_b, b_{\text{min}}, K), \text{ and} \]
\[ L = \sum_{i=1}^{n} v_{\text{len}}^i = \frac{t \cdot b_d}{b_b}. \]  

Therefore, individual sequence \( V \) consists of several segments \( s \), where each segment consists of several patterns \( v \):

\[ V = (s_{1,1}, s_{1,2}, \ldots, s_{1,n}, s_{2,1}, s_{2,2}, \ldots, s_{t,s}), \quad s_{i,j} = (v_1, v_2, \ldots, v_n), \quad n \in \mathbb{N} \quad (9) \]

The algorithm sequentially fills the segments of the rhythmic sequence. For each of the segments, it selects rhythmic patterns \( v \) until the sum of their lengths is equal to the length of the segment.

The distributions of individual meters and subdivisions are also governed by the individual levels. For example, level 11 contains simple rhythmic patterns that have an individual note length of a quarter (therefore, \( B_s \) with \( b_b = 4, b_{\text{min}} = 11, \) and any \( b_d \in \mathbb{N} \)). Level 24 contains more rhythmic patterns, which mostly contain pauses and are suitable for generating rhythmic sequences in meters \( 3/8, 6/8 \) or \( 9/8 \), where the basic period is eighth note.

### 3.4. Basic Gamification Elements

In order to increase the students’ motivation to use the application, we have enriched it with the basic elements of gamification. We use three gamified elements that relate to the students’ performance: a user profile and avatar creation, achievement badges, progression levels, and a leaderboard. The progression levels (Figure 4a) and achievement badges (Figure 4b) are visible on the student’s home screen. Students increase their level of proficiency by solving more exercises and progressing through levels of difficulty. The levels were defined by the teachers, and they vary from local orchestra to different competitions and international institutions. The badges, as shown in Figure 4b, reflect three different aspects of the student’s progress. The first aspect is accuracy: completing an exercise with(out) a certain amount of mistakes (from 50% up to 100% correct answers). The second aspect is the continuity of the student’s engagement with the platform: playing an exercise for a certain amount of days in a row—three days, five days, a week, two weeks, a month. The third aspect is the student’s speed: the amount of time needed to complete an exercise in five minute intervals, ranging from 25 min to five minutes.

In their profile, the students can also set or change their profile picture, username, institution, and school year (Figure 5a), with a visual representation of the collected points and achieved levels, and observe the leaderboard (Figure 5b). The leaderboard shows the cumulative points that were collected by individual players throughout their application use. More points can be achieved by both increasing the number of solved exercises and increasing the exercises’ difficulty. The students can observe their performance and compare it to the other players. By clicking on one of the platform’s users, the selected user’s profile page is displayed, with their achieved levels and badges.
Figure 4. Gamification elements. The left screen shows progression between multiple badge levels of proficiency, while the right screen shows the badges obtained during the practice.

Figure 5. The user profile and the leaderboard screens. The user can set their avatar and change their profile information (left). The user can observe the collected points of other players on the leaderboard screen (right).
3.4.1. Calculation of Points

While using the application, the student earns points by solving the exercises, which directly affects their achievements. Each exercise consists of two sequences that can be answered multiple times. After each completed exercise, the student’s points are calculated, being measured as a function of several factors: exercise difficulty (across the 16 difficulty levels), time taken (in minutes), number of corrections (additions and deletions of notes), the use of the metronome (yes/no), the number of submission attempts (checks for correctness), and whether the final sequence was correct. The sum of points can be either positive or negative, as shown in Table 1. The calculation takes the difficulty of the task, the time taken to answer questions, the number of notes added and deleted (penalizes random trying), and whether the user answered the question correctly (with a limited number of attempts) into account.

The difficulty levels gradually progress with a factor of 0.25 per level. The time factor significantly rewards the fastest players and more mildly penalises the slower players. The additions and deletions represent the number of each action performed by the user. The use of the aids, such as metronome and the number of checks for correctness, also mildly penalise the score.

<table>
<thead>
<tr>
<th>Table 1. Factor types and ranges for calculating the individual exercise score.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor Type</strong></td>
</tr>
<tr>
<td>Exercise level</td>
</tr>
<tr>
<td>Time</td>
</tr>
<tr>
<td>Additions Deletions $\delta$</td>
</tr>
<tr>
<td>Success</td>
</tr>
<tr>
<td>Used Metronome</td>
</tr>
<tr>
<td>Number of checks (before submission)</td>
</tr>
</tbody>
</table>

3.4.2. Exercise Automation and Instant Feedback

The automation of exercises generation enables the application to instantly assess the user’s response. The user’s response is assessed by comparing their response with the generated exercise while using several factors, as shown Table 1. The assessment is non-binary, thus motivating users who provided a partially correct answer, yet penalising the incorrect part of their response.

The automatic assessment is then used for instant user feedback. Besides the automatically calculated achieved points, the user is shown the correct response and their response below in two scrollable staves. The incorrect parts of their response (notes and other elements) are marked in red color. Therefore, the user can more closely observe and analyze the mistakes in their response.

The instant feedback element has multiple positive aspects to the learning process. First, it supports cause-and-effect learning, which is more difficult if feedback is delayed, as in the case of conventional pen-and-paper exercise submissions. Second, the automation of the feedback lowers the teachers’ work of assessing the exercises. Therefore, the computer assessment is performed in a non-binary evaluation through factors, which are important in ear training.
4. Experiment

The primary goal of the developed application was to provide an open platform, which would engage students and increase their performance in rhythmic dictation tasks. The application was tailored to increase student engagement through gamification elements and an intuitive interface on mobile devices. In our experiment, we wanted to assess whether these goals were achieved.

We evaluated several hypotheses. First, we assumed the mobile-friendly interface, specifically developed for rhythm input on a mobile device, will aid the students’ experience with the application. We gathered the students’ feedback through questionnaires. Regarding the automatic exercise generation algorithm, we also asked the students about their experience with the exercise difficulty. With the algorithm that was specifically crafted to aid a gradual learning curve and provide exercises, which are demanding, but not too difficult, we hypothesized the negative feedback on the exercise difficulty can significantly impact the students’ engagement.

Next, while the students might spend more time with an individual exercise at the beginning to get used to the interface, the time spent for solving an exercise should in time decrease due to familiarity with the interface and students’ increased proficiency. Therefore, we hypothesized that students’ time spent per exercise will decrease while solving more exercises within a single level.

The rhythmic dictation application was built to support rhythmic memory training, which is an integral part of the music theory learning process across the educational levels from elementary schools and conservatories to academia. During the development, we collaborated with the Conservatory of Music and Ballet Ljubljana, Slovenia. The application was primarily targeting the music theory classes, where it was also incorporated in the curriculum as a supplementary tool for rhythmic memory practice. During the development, we interacted with four students, who represented a sample of our target audience. We continually evaluated the students’ interaction with the application: whether they understood the user interface, whether the exercises were appropriately demanding, and whether the exercises were sufficiently interesting and engaging.

We then evaluated the application with the first and second year students at the conservatory. The students were randomly divided into one test and one control group in each year. The evaluation lasted five weeks, during which we held four in-class meetings with the students of the test groups. The students were asked to use the application during the meetings through a group student challenge, during which the students competed to achieve points in the application.

After the five week period, the students of both test and control groups participated in a standard curriculum exam. We compared the exam results and observed the application’s impact on exam performance. The test groups consisted of 11 first and 12 second year students, while the control groups consisted of 11 first and 13 second year students.

In this section, we first describe the student challenge, followed by an analysis of the collected questionnaire responses from students. Finally, we describe the students’ results in a conventional exam and evaluate our hypotheses regarding the application’s impact on the students’ performance.

4.1. Application Evaluation—Student Challenge

During the five-week application evaluation period, four meetings with the test group students were organized. The meetings were held during the music theory classes. Our goal was to observe the students’ engagement with the application. We proposed a student challenge, where the students competed to gain points and rank high on the leaderboard, to gain the interest of students.

At the first meeting, students from both control and test groups were present. At the second meeting (second week), the platform and rhythmic dictation application were presented to the test groups. We presented the application evaluation to the students as a month-long challenge. The students were also asked to actively use the application in-class. At the third meeting (week 3), the test groups students again used the application during the music theory class. The fourth meeting (week 5) was organized to conclude the challenge and nominate the winners of the challenge.
Throughout the meetings, we investigated the application’s use and students’ experience through a series of questionnaires.

4.1.1. Test Group Evaluation Schedule

At the first meeting, the students of the test groups were given a questionnaire that contained general questions regarding the use of tools for practicing music theory on mobile devices. The first part of the questionnaire involved questions about which applications (including music theory apps) the students use on their mobile devices. The second part of the questionnaire consisted of questions regarding the students’ rhythm practicing at home.

During weeks 2 and 3, we enabled access to the application to the test groups. We conducted a live challenge during a music theory class. The goal of the challenge was to increase student engagement with the application, by enticing them to gather points and rank high on the leaderboard. During the challenge, we motivated the students further by presenting intermediate results live on a classroom display. Symbolic rewards were given to the first three students. During the week 3 meeting, we also distributed questionnaires to the test group students. We asked the students about their experience with the application. The responses to the questionnaire were mostly positive (as discussed further in Section 4.1.2).

Five weeks after the beginning of the application’s evaluation period, we organized the fourth meeting. We presented the final results of the participating students and handed out plaques to the winners of the challenge. All of the students received symbolic rewards in gratitude.

4.1.2. Questionnaire Results

In the initial questionnaire, given during the first meeting, all of the students reported using mobile applications, such as social, messaging, and music apps (SnapChat, Instagram, FB Messenger, YouTube). 79% (14 students) of the students reported using mobile apps for learning new skills, such as foreign languages and instruments. However, applications for practicing music theory, such as Teoria.com, TonedEar and MyEarTrainer, were rarely mentioned. Only a few students (17%, three students) used various rhythmic dictation exercises. In the second part of the questionnaire, most of the students reported practicing rhythmic exercises at home (67%, 12 students), which was expected, since the students had no option at home but to read and perform the rhythmic exercise in a conventional way of listening to pre-recorded examples. However, they showed mixed opinions on whether they wanted additional ways to exercise rhythmic dictation, as 55% (10 students) of students did not want additional rhythmic dictation exercises. Therefore, the shift in their opinion posed a key challenge for the success of the proposed application.

During the third week, a second set of questionnaires was given to the students. The students were asked whether the application was difficult to use—all of the students answered negatively (100%, 18 students). Most of them answered that the exercises were not too difficult (16 students, 89%) and that they got used to the application’s use over time (16 students, 89%). The answers were consistent with the goal that the application exercises should not be too difficult. Most of the students responded that the rhythm input keyboard worked as intended (13 students, 72%). The majority did not use additional paper and pencil (17 students, 94%) exercises. However, the sound of the organ that was used by the application was perceived as disturbing (16 students, 89%). Many answered that they would rather listen to the piano because they are more accustomed to it. Two-thirds of the students had enough time to complete the exercises (12 students, 67%), and did not adjust the speed of the dictation playback (67%, 12 students). The students also positively responded to the live challenge. The live comparison and symbolic rewards have proven to be an extremely important motivating factor. While the students used the application during the live challenge, they were not limited from, or required to use the application at home. In the usage frequency visualization, as shown in Figure 6, the increased use after the individual meetings shows the students began using the application at home more often after the second meeting, as compared to the first week.
During the final meeting (week 5), we again asked the test group students to respond again to the questionnaire that was handed out during the third meeting. We investigated the changes in their opinions after one month of application use. Again, we received positive responses. The students replied that the exercises were not difficult (91%, 16 students). All of the students got used to the application during this time. To most students, the rhythm input keyboard functionality seemed logical and worked as intended (83%, 15 students). All of the students began using the application to practice and stopped using the conventional paper and pencil practice. The results were somehow expected, since the students were well acquainted with modern technology and were no longer as paper-bound as older generations. The students also grew accustomed to the sound of the organ that was used for playback (72%, 13 students—Figure 7).

![Graph of the application’s usage frequency per day. The second and third meetings are marked individually for each test group. The engagement the day after the meeting (# of exercises) was initially low after the meeting with both the first- and the second-year students. However, it increased after the second meeting.](image)

Most of the students had enough time to complete the exercises (91%, 16 students), and did not adjust the playback speed (91%, 16 students). When asked about whether they showed the application to their friends, the majority responded positively (73%, 13 students). In their final remarks, the students highlighted the following features:

- the students liked the ability of using the application on a personal computer in addition to the mobile device,
- the scoring and achievements (badges) were motivating, and
- the ability to pause/stop the playback was helpful.

### 4.2. Analysis of the Exercises’ Data

Twenty-three students, 11 from the first and 12 from the second year, completed 496 exercises in total (Figure 8). Each exercise consisted of two rhythmic sequences that could be answered multiple times. In total, the students correctly answered 837 sequences. The first-year students averaged 24.5 sequences and the second-year 26 sequences.

The rhythmic dictation application is organized into 16 difficulty levels. The student had to complete at least 12 exercises at the current level in order to advance to a higher level. When starting the application, the student could choose which level to start at from the subset of levels they already achieved. 39% (seven students) of the students remained at the first level (level 11), because they did not complete enough exercises to pass onto the next level, while others moved to higher levels. During the evaluation, only one student reached all of the rhythmic levels available.
Figure 7. Initial responses to individual questions about the students’ behaviour and experience while using the application, and the responses after one month of application’s use. The most significant change is in the students’ perception of the organ sound. No student reported still using pen/paper while using the rhythmic application.

Figure 8. The number of responses (sequences solved) per levels (left), and the average number of seconds spent per exercise in first ten consecutive exercises for first three levels.

We also observed the time needed to complete the individual exercises. In their first exercises, the students needed more time than in later repetitions (Figure 8)—the average time gradually decreased with the number of exercises played. With increasing difficulty levels, we noticed that the number of event deletions increased. The number of dictation plays remained steady across all levels of difficulty. The number of attempts to solve also remained steady, with the exception of level 11 (1/16 difficulty level), where the sequences were trivial for the conservatory-level students to solve. Figure 9 shows a subset of players who played at the most common levels.
Figure 9. The frequency of application’s use, sorted by the final achieved level of individual user. The circle color represents individual levels (11, 12, 13, 22, and 24), while its size shows the number of games played by the individual user.

The gathered data confirmed our assumption that some of the observed values, such as time spent, decreased over time, while others remained steady due to the increasing difficulty of exercises. The students’ engagement in out-of-class use gradually increased, which we consider to be a confirmation of the application’s usefulness—students liked the interface and found it easy to use—as well as gamification elements, which, through the student challenge, brought competitiveness into interaction between students.

4.3. Exam Performance

We also analyzed the results of both control and test student groups by observing their performance in a conventional in-class exam in order to confirm our hypothesis about the experiment’s impact on the students’ performance. The exam results were graded using grades 1 (worst grade) to 5 (best grade). The first-year control group students achieved an average grade of 4.3, while the test group students achieved an average of 4.5 (4% increase). The results were statistically tested while using the Mann–Whitney U-test and the difference was not statistically significant (U = 16, p > 0.05). A larger difference was observed for second-year students, where the control group achieved an average score of 3.58 and the test group 4.44 (19% better, significant difference, U = 24, p < 0.001). We can confirm that the use of the application had a positive effect on their performance in the exam, as better results were achieved by students using the rhythmic dictation application. As both of the
groups were relatively small, we also used a resampling method to compare the group averages. At 1000 replicates, the method estimated a 69.2% probability that the average test group score was greater than the average control group score for the first-year student groups. For second-year results, the algorithm estimated this probability at 99.6% at 1000 iterations. These estimates confirmed the Mann–Whitney U test; therefore, showing that students who used the application performed better in the exam.

The difference between the first (no significant impact) and second year (significant impact) student groups could be attributed to several factors. While the five-week challenge in combination with the application’s use produced an improvement in the students’ results, our approach could have a less motivating affect on the first-year students, as compared to the second-years. Moreover, the second-year students have undergone a larger amount of training at the conservatory level, possibly resulting in higher self-discipline for training. In addition, rhythmic dictation and ear training is itself a specifically frustrating task for some students. Given the fact that the first-year students come with various levels of knowledge and have yet to overcome their personal frustrations with certain aspects of training, this could indicate that the application has not been a substantial aid to the first-year students. The students who ranked highest in the competition rankings were also from the second year test group. However, both student groups were small and a larger longitudinal study is needed in order to further confirm the results of this evaluation, and fully evaluate the application’s impact on the learning process and performance.

Based on our previous work, where we also observed the positive impact on the students’ performance while testing the intervals dictation application without the five-week challenge, we assume that the challenge is not a substantial contributor to the increase in the students’ performance. However, further research is planned to identify the underlying causes for this difference in the increase between both years. Because we included all conservatory students enrolled in first and second year—either in control or in test groups—as further confirmation that the evaluation results could also be achieved by performing a similar experiment on more institutions, in order to observe the application’s impact and confirm, whether the results gathered within this study could be generalised.

5. Conclusions

In this paper, we presented the rhythmic dictation application, implemented into the Troubadour e-learning platform. The platform is an open-source, web-based platform for music theory learning. The platform was developed in order to engage students in music theory learning by offering them a flexible and individualised medium for practice through carefully tailored exercises. The platform is open source and it is publicly available on Bitbucket. It offers a supportive and flexible environment for the teachers and students. The rhythmic dictation application features a mobile-friendly user interface supported by gamification elements for attaining the students’ engagement, while offering a flexible environment for the teachers. We investigated several aspects of the application—the user interface, the automatic exercise generation algorithm, the students’ engagement, and their exam performance. The user interface and its keyboard for rhythm input was well-accepted. The automatic rhythmic sequence generation algorithm was evaluated through students’ feedback regarding the exercise difficulty and their engagement. To further engage students, we created a five week challenge during which the students were asked to use the application through a gamified experience of collecting points and badges, which were visible to other students. In order to confirm the hypothesis about the application’s impact on the students’ performance, we compared the conventional exam results of the test groups students, who used the rhythmic dictation application, to control groups of students, who did not use the application. The evaluation showed that students support the use of a gamified rhythmic dictation application. The students also reported a very positive experience with the application, which was further substantiated by the claim that most of them would recommend the application to their friends. Therefore, the students supported the idea of an application, which offers them a modern medium of learning music theory as an additional (non-conventional) tool for ear
training and music theory learning. The developed user interface for rhythmic dictation was considered easy to use. Nevertheless, there are several aspects that will be further improved in our future work through a deeper user experience analysis and iterative development of the interface.

The results of the A/B test, in which we assessed the effect of rhythmic dictation application’s use on the students’ performance in exams were also promising. The comparison of exam results between the control and test groups showed a positive impact of the application’s use on exam results, which was statistically significant for the second year test student group. Even though all first- and second-year students at the Conservatory were included in this study (either in control or in test groups), the results suffer from a relatively small student groups. The study was carried out at the Conservatory of music and ballet, Ljubljana, Slovenia, which represents roughly 50% of the state-wide student population enrolled in a music programme at the conservatory level. Based on the evaluation presented in this paper, we can corroborate that the gamified rhythmic dictation application aids the students’ performance, which we attribute to gamification and automatization of rhythmic dictation exercises. In addition, the results of the conventional exam are a consequence of the application’s use and the five-week challenge we underwent with the student test groups. Although our previous work [31] shows the beneficial impact of the platform on students’ performance, the experimental setup in this paper includes in-class time spent with the students through the challenge, which could also be a contributing factor to their increase in performance. Finally, the results of this experiment should not be too quickly generalised onto more general groups, such as elementary music-school students. The students at the conservatory are above-average skilled, as compared to lower-music school students.

6. Future Work

The application and platform currently only include basic gamification elements. Our future work includes several more advanced aspects. First, we plan on including introduction tutorials to aid student onboarding, which was done in-class for the experiment described in this paper. When considering the late shift towards e-learning, this option has become less possible. Second, to aid students with different levels of knowledge, i.e., first-year students, we plan on further exploring the option of personalised scenarios and personalised automated exercise generation to ease the learning curve, retain engagement, and consequently improve performance.

Consequently, the exercise scoring system presented in this paper could be further simplified and made transparent to the user. In its current version, the application shows the achieved score only at the end of the exercise. Therefore, the users familiarize themselves with the scoring system through the inference with the exercises. We plan on adding the continuous score label to the interface, which would inform the user. One of the options, which is commonly used in games, is to show the maximal possible score, which can be achieved, and lower the score accordingly with each retry and time spent solving the individual exercise. This way, the users will be potentially more aware of how their actions influence the score; for example, by waiting too long to complete the exercise, or just trying out their answers many times.

Based on the results that we collected in this study, we observed significantly better performance in the final exam by second-year test group students, as compared to the control group. Our current and future work is focused on further development of applications and other learning tools, and evaluating their impact on the students’ performance. When considering the current shift towards remote learning due to Covid-19, the Troubadour platform offers a unique learning tool for ear training. We plan on further extending its capabilities by including a new application for chord progression training, and real-time remote interaction between the teachers and the students, in order to support “remote in-class” interaction.
Author Contributions: Conceptualization, M.P., PŠ. and M.M.; methodology, PŠ. and M.M.; software, L.S. and M.P.; validation, L.S. and M.P.; formal analysis, L.S. and M.P.; investigation, L.S. and PŠ.; resources, M.M.; data curation, M.P.; writing—original draft preparation, M.P. and L.S.; writing—review and editing, M.P., M.M. and PŠ.; visualization, L.S.; supervision, M.M.; project administration, M.P.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

9. Wiggins, B.E. An Overview and Study on the Use of Games, Simulations, and Gamification in Higher Education. *Int. J. Game-Based Learn.* 2016, 6, 18–29. [CrossRef]


