Rhythm Rangers: an evaluation of beat synchronisation skills and musical confidence through multiplayer gamification influence

Francisco Rosado Correia
Sound and Music Computing
Aalborg University Copenhagen
Copenhagen, Denmark
frosad19@student.aau.dk

Marco González Pérez
Sound and Music Computing
Aalborg University Copenhagen
Copenhagen, Denmark
mgonza19@student.aau.dk

Ramon Romeu
Sound and Music Computing
Aalborg University Copenhagen
Copenhagen, Denmark
rromeu19@student.aau.dk

Rasmus Kjærbo
Sound and Music Computing
Aalborg University Copenhagen
Copenhagen, Denmark
rkjarb19@student.aau.dk

Vatsal Guruvayurappan
Sound and Music Computing
Aalborg University Copenhagen
Copenhagen, Denmark
vguruv19@student.aau.dk

***Abstract*—Participants' musical confidence and beat synchronisation skills are explored before and after playing the custom made multiplayer game, Rhythm Rangers. The effect on timing variability after short-term practice (priming with a baseline test) is investigated and correlated with a pre- and post-experiment self-evaluation questionnaire focused on musical sophistication, behaviours, and confidence. The baseline test is repeated after participants have played the game in order to measure the training effect. For both the test and game, participants clap along to the beat of either two or four different audio loops ranging from a quarter-note metronome, simple drum loop, syncopated drum loop, to a complex break beat. In order to conduct the experiment and play the game, custom made wearable controllers were built and the necessary software developed integrating multiple ESP32 TT-GO T-Audio micro-controller boards, the microcomputer Raspberry Pi 3, and Pure Data creating both the multiplayer game and an environment for psychological experimentation. Baseline test results indicate better beat synchronisation when stimuli is a drum loop compared to a metronome. The same was true for game scores, where the average standard deviation of clap scores was highest for the metronome sound. A general downward trend for average drift variability (calculated as the standard deviation obtained from total local data scores) was found for both test stimuli (quarter-note metronome and simple drum loop). The total average standard deviation decreased with relation to the amount of information in the audio loops until a certain point where we assume the perceived complexity of the audio loop got too complex compared to the players' skill level. Despite our efforts, little correlation was found between the qualitative data and the participant performance during the experiment. All code is available on Github[1].**

***Index Terms*—Beat Perception, Beat synchronization, Gamification, Clapping Task, Timing Variability, Rhythm Training, Musical Confidence**

[1]http://bit.ly/RHYTHM-RANGERS-SMC07

I. Introduction

Gamification is a wonderful way to promote learning and skill development [1] [2]. This powerful tool creates incentives for progression and motivates skill development [3]. Rhythmic coordination is an essential part of everyday life (e.g. walking) and it is involved in higher order cognitive tasks like dancing and performing music. Research on the coordination between perception and action has seen a surge in the last 20 years, especially when it comes to music [4]. A great deal of classic laboratory studies on perception of isochrony in sensorymotor synchronization (SMS) often use a simple metronome of pure tones or clicks as stimuli. The participant finger taps along with the tactus while the researcher records the performance. At some point, the stimulus is left out and the participant continues tapping following their own internal 'metronome' [5] [6]. Newer studies incorporate various forms of both movement (tapping or exciting various measurement devices with or without feedback, breaking a laser beam, using limb extensions to trigger drum pads, etc.), modality of stimuli (ranging from simple drum loops to excerpts of full compositions, or using visual stimuli) to the form of coordination; either on-beat or off-beat, in-phase [down beat synchronization] or inverted-phase [up-beat synchronization] [7] [6]. Modern approaches to experimental music psychology often incorporate parts or whole batteries of sophisticated tests (i.e. Seashore, Gold-MSI, BAASTA, H-BAT, BAT, Musical Ear Test, etc.) while others develop their own tests on a per-experiment basis. The following experiment draws inspiration from classic experimental psychology for rhythm tapping [5] and couples it with modern approaches to beat synchronization [8] and gamification [1] [3].

In reviewing beat synchronization literature, we found that optimal beat reproduction (e.g. tapping to the beat) is obtained with an interonset interval (IOI) around 600 ms (100 BPM). According to a study performed by Ehrlé et al., (2005), minimal variability was observed with IOIs between 200 ms and 1200 ms [6]. Ehrlé et al. (2005) reference a study by Getty (1976) [9] where they found "that variance increases linearly with interval durations between 300 and 900 ms attributing this effect to Weber's law" [6]. According to Weber's law of 1979, the just noticeable difference in stimuli between two tempi is linear [6] - this was found to be a good general conclusion for tempi in ranges between 400 and 800 ms according to a study by Friberg and Sundberg (1995) [10]. Based on this, the two game tempos; 90 BPM (IOI = 666.67 ms) and 100 BPM (IOI = 600.00 ms) were chosen and the game was constructed to follow the gestalt principles; law of proximity (equal spaced rhythm divisions for the clapping task), continuity (each player should avoid clapping exactly four times per game in randomised order, but otherwise clap), and closure (each game lasts exactly 32 beats regardless of performance). The aim of this project is to evaluate participants' beat synchronization skills and musical confidence through rhythm training with a multiplayer game. To achieve this, a baseline skill evaluation test was constructed along with a four-player rhythm game in which the players have to synchronize their clapping to different 2-bar audio looped 4 times (total of 8 bars) of varying rhythmic complexity (from quarter-note metronome to a jungle break beat). The game doubles as a psychological experiment, as the player device records the action and movement of the player. The quantitative pre-game and post-game baseline tests along with game data from the devices are correlated with a qualitative pre-game and post-game questionnaire in order to evaluate the effect of the game and experiment as a whole. Players enter the game by strapping the wireless custom made game controller to the back of their hand (see Figure 4) and follow the game instructions. Four players then try their best to synchronize their claps to four audio loops with increasing complexity - this at the tactus / quarter-note level. Each game has a total of eight trials; four audio loops at tempo 90 BPM (666.67 ms) and four similar audio loops time-stretched to tempo 100 BPM (600.00 ms). According to a study by Begel et al. (2018) "there is a tight link between rhythmic skills, motor, and cognitive functions" [1]. The same study lists an impressive corpus of research suggesting the positive affect rhythmic skills might have on motor and cognitive functions. Treatment and training of people suffering from neurodegenerative disorders like Parkinson's disease with rhythm and music have shown to positively affect traits like gait, stride length and walking speed when pacing along with music or a metronome [1]. Interestingly, the "positive response to rhythmic stimulation depends on patient's perceptual and sensorimotor rhythmic skills, thus pointing to a strong link between rhythm processing and motor control" [1]. End on our own words here.

A. *State-of-the-art*

A brief description of related projects:

1) Rhythm games: Rhythm Workers was the final product of a research project about how rhythm training could improve cognitive functions of people having neurological disorders [1]. This game uses rhythmic patterns and musical stimuli as a tool to trigger those cognitive abilities. Furthermore, the game is composed of two versions, a perception and a tapping version, in order to evaluate the training of the players.

In the perception version of the game, the training is carried out using an adjusted version of the Beat Alignment Test [8]. The main goal of the game is to construct a building and in order to achieve that, the player has to detect whether a metronome is aligned to the beat of the stimulus or not. In the tapping version, the goal is to tap to the beat of the stimulus as accurately as possible. If the player taps synchronously to the beat (tapping version) or if they correctly detected whether the percussion sounds were aligned to the beat (perception version), the building would appear better structured and more aesthetically appealing than when the player's performance was not good.

Beat Saber is another rhythm game implemented in VR where the goal is to "cut" the beats with hand movements (shown as small cubes) as they appear on the display. Every beat indicates which hand you need to use and also the direction you need to match. Given this, the players almost dance while cutting the cubes and avoiding obstacles. Each cut is strongly supported by sound and visual effects to emphasize the rhythm. This game is a good example on how body movements can be used on a rhythm game in order to make the experience more enjoyable for the players. In addition, the sound and visual cues displayed for the player to emphasize the beats were also a big inspiration for the implementation of our game[2].

2) Wearable devices: As this project aimed to make use of the player's body movements, more precisely arm movements, some research was carried out during the beginning stages of the casing design. The following devices were found: Soundbrenner Pulse is a smart vibrating metronome that allows the user to follow the beat without hearing sound, but feeling the vibrations on your chest, wrist, or ankles. In addition, it works hand in hand with a metronome mobile app[3].

GripBeats is a device created with the ambition of allowing non-musicians to make music but also professionals to explore a different way of making music. It is shaped as a bracelet that straps around the wrist and extends to the hand[4].

The mentioned devices aided in the decision of which path to take for the design of this project's wearable game device.

B. *Synchronisation: Variability per beat (local score) and overall per trial (drift score)*

A study performed by Madison et al. (2013) [7] investigated isochronous serial interval production (ISIP, classic continuation tapping) by measuring the difference between individual intervals respectively (per beat), the local variability,

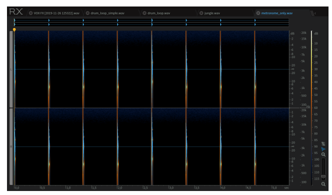
[2]https://beatsaber.com
[3]https://www.soundbrenner.com
[4]https://www.gripbeats.com

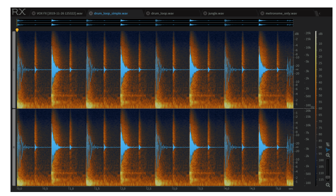
and by looking at changes in long-term variability, the drift. According to the authors, analysis of the variance can be further elucidated by decomposing the ISIP task into local variability - which could be reflecting random neural noise - and the slower evolving drift variability - which could reflect some aspect of the participants short-term memory. During the first hour of practice, their participants showed a substantial decrease in variability, but not much afterwards. For different mode of response, amount of feedback, and interval duration, they found similar effects; a reduction in both local and drift variability. According to the authors: ”the results suggest mainly effects on motor implementation rather than on cognitive timing processes” [7].

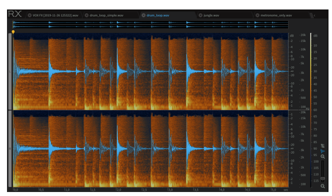
Madison et al. (2013) suggest that the two different measures of ISIP variability (local and drift) might not be affected equally by training [7]. They present a glorious corpus of research which suggest that the two measures are affected differently when it comes to state motivation, sensory feedback, BPM / IOI, reaction time variability, and cognitive load [7]. Distinctions between the variance associated with local and drift variability is not clear from the literature though. According to Holm et al. (2013): ”On the one hand, drift variability is more affected by sensory feedback than is the local variability component and the local component is more strongly related to intelligence. On the other hand, drift variability is arguably more dependent on short-term memory and should thus be more susceptible to cognitive load” [11]. An interesting suggestion by Holm et al. (2013) decouples a direct link between executive functions with motor timing, unless the executive functions are required in order to control complex motor coordination. In that case, executive functions can indirectly affect the timing performance [11]. It seems the literature supports the hypothesis that ”motor coordination is the most important factor behind the interference effects found in interval timing” [11].

C. *Gamification, multiplayer, and flow*

Csikszentmihályi [12] lists various conditions for optimal flow experience; e.g. the task should be balanced in regard of difficulty and participant skill level and there should be instant feedback on how well you are performing the task. When the skill-to-task balance is optimal the player can experience a merging of action and awareness (e.g. the feeling of events happening naturally). By coupling clear and concise goals with immediate visual feedback, we hope to grab the players' concentration on the task at hand in order to achieve flow states like transformation of time (immersive engagement) and loss of self-consciousness (decrease of negative personal assumptions towards musicality). If the task is too easy the participants might become disengaged. If the task is too hard, apathy might occur [13]. In a review by Begel et al. (2017) [3] 27 rhythm-based games available on the market was reviewed. According to a rhythm training protocol set by the authors, none of the reviewed games live up to the necessary criteria of the protocol (see [3] for the complete list of reviewed games), albeit presenting good grounds for rhythm training. Mostly, the poor precision of temporal movement recordings made motor performance variability (drift) hard to measure along with testing the precision of beat synchronization (local) [3]. Another important factor is how the reviewed games handle leveling and difficulty increases. Guitar Hero for example increases the difficulty by adding more events that you need to respond to, this without affecting the auditory stimulus (you play along to the same song). An important aspect of the rhythm training protocol by Begel et al. (2017) is varying beat saliency as a control for levels and difficulty. Levels and difficulty handling in Rhythm Rangers varies beat saliency by presenting loops of varying complexity: starting with only a metronome sound, then changing to a simple drum loop with kick drum, snare drum and hihat sound only on straight beats (eight-note level). The difficulty increases with the next drum loop which has added syncopated subdivisions in it. The final drum loop is an uptempo breakbeat with many syncopated subdivisions (dubbed the ”jungle” drum loop) see figure 1.


(a) FFT metronome


(b) FFT simple drum loop


(c) FFT syncopated drum loop

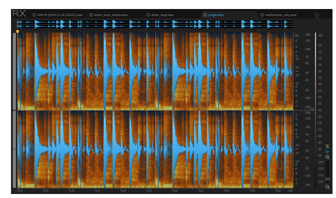

(d) FFT jungle breakbeat

Fig. 1: Spectrograms of the different sound samples used for the experiment.

D. *Aims and Predictions*

Within the scope of this project, it is hoped that players who play our game can improve their synchronisation skills in a multiplayer game, a fun yet challenging medium. This improvement is to be understood as improving the detection of the tactus while at the same time clapping as close as possible minimising local variability and drift variability. It is also hoped that participants with higher scores during the pre- and post-test and game phases of the experiment will feel an increase in their musical confidence and therefore are more likely to engage in future musical activities. It could be assumed that participants claiming they have more rhythmic experience (e.g. drumming, playing rhythm games, etc.) will have their abilities reflected in high scores during the game. Perhaps the prime objective for this project is to create a game that is fun and engaging.

E. Abbreviations

- VM - TTGO T-Audio board running the Visual Metronome
- RPi - Raspberry Pi
- Pd - Pure Data
- OSC - Open Sound Control
- UDP - User Datagram Protocol

II. Game device design & implementation

A. Designing and building the wearable device

1) The main case: The TTGO T-Audio board is attached to a wooden laser-cut base plate (see figure 2), using nuts, bolts and spacers. Using the Epilog Fusion M2 laser-cutting machine, holes were laser-cut into the base plate, according to the exact dimensions of the device, so it could be strapped and attached firmly onto the back of the player's hand. There also exists a sturdy foam layer, in which holes were cut to thread the elastic band through, and is pasted to the bottom of the base plate so the player has a comfortable device fit throughout the game-play. The use of nuts and bolts, mainly ensures a firm attachment between the device and base-plate, keeping the overall structure sturdy for the player during game-play. Furthermore, the spacers create a gap between the device and the base plate, which is sufficient to accommodate a 3.7V Li-Po battery. The elastic band is threaded through the base plate holes, and sewn at the ends, so the casing could be strapped onto the player's palm (see figure 4). The purpose behind sewing the ends of the elastic band is to ensure a 'one-size-fits-all' form of comfort and flexibility, along with a tight fit, which is available to the average hand size. The top surface of the device, containing the LED lights, is positioned to face away from the player so when claps are being performed during game-play, the LED lights are clearly visible showing the performance of the player in real-time. Figures 3 and 4 show the final case design.

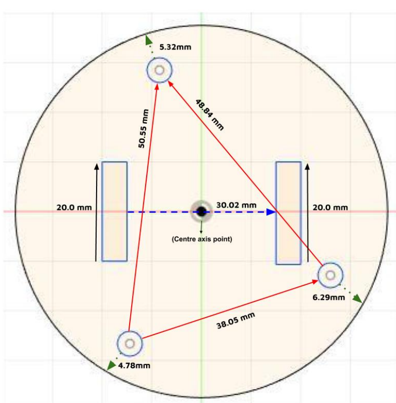


Fig. 2: 2D schematic of the wearable game device base plate.

2) Usability test feedback: An early prototype of the game player device was made in order to make a usability test. The feedback received pointed out a certain amount of wobbling of the device which occurred when clapping, especially with intense force. Also recommended was to improve the fitting and placement of the battery in the casing, and finally to make the elastic band more adjustable according to the desired level of tightness. The final recommendation, was a soft cushion or foam material layer between the hand and the device, so as to give a comfortable experience. The suggestions were implemented in the final design.

Fig. 3: Top view of final four game devices with player numbers and colours.

Fig. 4: Device with case, strapped to the player's hand.

B. Technical specifications and game implementation

1) Hardware and programming: The game was implemented using the microprocessor ESP32 on TTGO T-Audio v1.6 board[5]. The ESP32 is an affordable micro-controller providing built-in Bluetooth and WiFi support, several GPIOs and analog inputs. It is a powerful chip in terms of carrying out real-time DSP applications[6]. The chip is mounted on the TTGO T-Audio, that also has the MPU9250, a 3 axis acceleromenter and gyroscope. The board also provides 19 built-in RGB LEDs and a battery managment circuit. This kind of device seemed adequate for this type of project due to its shape and weight in order to analyse the movement data of each player. Four boards were used in the project as wearable devices for the players while one was used as a visual metronome (VM). In addition, the latter runs a web-server, allowing for the scores of each player to be displayed.

[5]https://github.com/LilyGO/TTGO-TAudio
[6]https://faust.grame.fr/doc/tutorials/index.html

Application Layer	Raspberry Pi	OSC/HTTP
Transport Layer	UDP	
Network Layer	IP	
Data Link and Physical Layer	Ethernet	WiFi

Fig. 5: The network layers of the project.

To establish communication between all devices, a WiFi network was created using a router in order to use the Open Sound Control (OSC) protocol over the User Datagram Protocol (UDP). The OSC protocol is supported by both the ESP32 and Pure Data (Pd). This brought the benefits of modern networking technology to our game and testing environment[7]. In addition, UDP was chosen instead of the Transport Control Protocol (TCP) due to the fact that it prioritises lower latency above reliable packet delivery due to the absence of handshaking. Since real-time interaction and communication was necessary for this project to succeed, it was decided to use UDP though it resulted in minor problems with loss of packets during transmission. This was most prominent in areas with high amounts of wireless network traffic. To mitigate, the WiFi channel with the least activity on was selected before each run of the experiment and devices capable of communication with a wire (RPi and laptop) were connected directly to the router. A RPi running a Pd patch (programme) was used to distribute the messages between ESP32 boards via the OSC protocol. Arduino IDE was used to program the ESP32 due to the fact that it is open source and some members of the group were already familiar with it. Once the complete setup was ready, two main programs were implemented: the one that is running on the wearable game device and the one loaded onto the VM.

2) Inter-device communication: A big challenge of this project has been to enable a reliable communication between all the TTGO T-Audio boards and the Pd patch on the RPi. Bluetooth was considered, but discarded as it has higher latency and shorter range than WiFi [14].

The most challenging part of this setup was being able to synchronize the VM with the wearable game devices. This synchronization needed to be as accurate as possible as the players need to follow the VM, so the wearable game devices should have the same tempo in order to correctly detect the local variability for each clap. After testing different methods, we decided to follow the procedure described in figure 6.

1) When the Game Start button (I/O 36) is pressed on the VM, it sends an OSC message to the RPi stating that the game is starting.
2) The RPi sends an OSC message with the BPM in milliseconds (e.g. 90 BPM would be 666.67 ms) to the VM and to the wearable game devices. Immediately afterwards, the RPi sends a metronome start message to the VM which blinks it's LEDs in time with the count-in before the game starts. The actual game starts after the count-in, on the following quantised beat, when the game start OSC message is sent to all the devices and the respective audio sample chosen for the test or game starts playing, the VM flashes accordingly and the game player devices starts detecting claps.
3) During the game, each wearable device sends the score for each clap to the RPi, which stores it for future data analysis. At the end of the game, each wearable game device sends the average (drift) score of all claps.
4) Once the game finishes, the RPi calculates the final score (with game penalties, mentioned in section 3C) from each player and sends it to the web server hosted by the VM. Final scores are then displayed on the web server, shown on a screen to every player.

[7]http://opensoundcontrol.org/introduction-osc

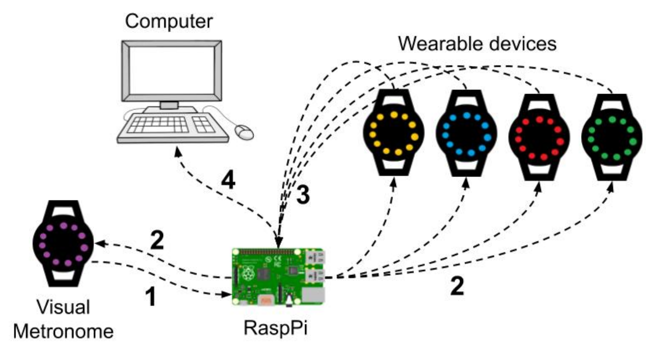


Fig. 6: Scheme of the communication between devices during a game.

Hard syncing was attempted by sending the game tempo clock from the RPi to all devices. With this method it was observed that there was a noticeable and odd desynchronisation between the devices so this method was ultimately discarded.

3) Visual metronome routine: The TTGO T-Audio board acting as a metronome provides a visual aid to follow the tempo by synchronising the LEDs blinking rate to the rate of the audio metronome. The VM will blink for the duration of 8 bars, then the game stops. The metronome can be stopped by pressing the start button while it is running. There is also a mode button (I/O 39) which switches between the different sections of the experiment. The VM and the Pd patch both start off in the pre-intervention baseline test mode, in which each individual participant listens to the audio loops and claps along to the beat (more on section 3B). Once each participant has completed the baseline test, the mode button is pressed and the VM changes to game mode. In this mode, the VM displays different coloured lights pseudo-randomly dictating the rules of the game. During the game, the VM displays either a pink light in sync with the audio metronome, one of the player colours or a white light, depicting the rules of the game (see section 3C). Each game has a total of two white lights and two of each player colours' lights displayed and their order is randomised per game as shown in figure 7. The mode button is pressed one more time and the post-game baseline test mode is loaded. Pressing again brings the VM back to pre-intervention baseline test mode, essentially allowing the cycling of the modes.

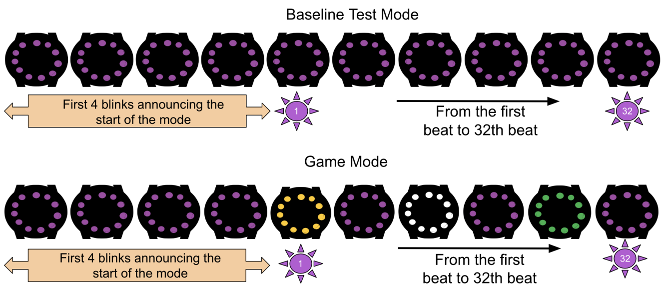


Fig. 7: The VM LED behaviour for baseline test and game modes.

A web-server was implemented into the VM so that the score for each player can be displayed in a web browser on the static IP address of the web server (http://192.168.1.50). After each round, it is necessary to refresh the web page to display the latest scores.

4) Wearable device routine: The purpose for the wearable device is very straightforward: to detect claps and to determine the accuracy of these claps in respect to the tempo. On successful boot (connection to the WiFi established and code initialised), each device then displays the respective player colour on the LEDs and waits for the game start OSC message. The game device constantly reads the acceleration values captured by the accelerometer. When a clap is detected the score is calculated depending on how close or far it was from the beat using the following expression:

$$Score = (1 - \frac{abs(x)}{y}) \cdot 100 \tag{1}$$

where x represents how early or late the player clapped in relation to the tactus in milliseconds and y is the BPM in milliseconds. The clap scores are normalised and thereby made independent of the tempo so that a clap score value is obtained as an integer between 0 and 100; where a score of 100 is equivalent to a perfect clap on beat, and 0 is a perfect off-beat clap. After that, the LEDs on the board light up showing the clap score: the higher the score, the greener the lights will light up, and the lower the score the more red they will be. The score is then sent to the RPi. Once the game finishes, the average of all clap scores is calculated and sent to the RPi. The player game device then reverts back to its initial state, waiting to start another game.

5) Clap detection algorithm: An important part of the project was being able to distinguish claps from non-claps (regular motion) and the ability to log missing claps in the Pd patch. The accelerometer values obtained are used to detect sudden changes in acceleration on all axes. When a clap happens, there is a high peak of acceleration when both hands make contact. Since people clap in different ways, it was decided to calculate an overall acceleration value for all axis using the following equation:

$$a_{sqrt} = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{2}$$

where a_x, a_y and a_z are the acceleration values per axis. By doing this, high peaks are detected in any of the axes, and will always be a positive number.

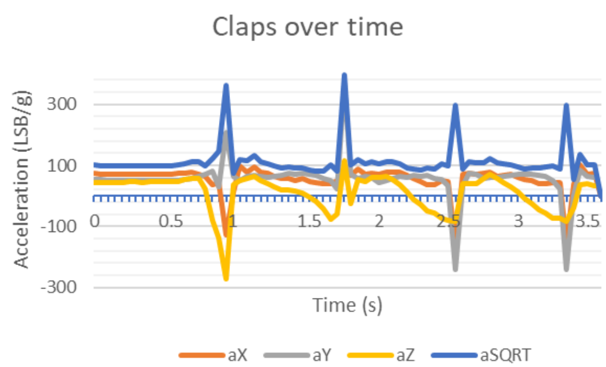


Fig. 8: Acceleration data from the MPU9250 accelerometer detecting claps.

As seen in figure 8, the acceleration peaks for the a_{SQRT} are very prominent when a clap occurs, so it was decided to set a threshold. When the acceleration value surpasses that of the threshold value, it establishes that a clap is detected. Some false and wrong results were still being obtained with this algorithm so inspired by Cyril Arnould in [15], it was decided to average the last five acceleration values and compare it with the current acceleration value, and then, if the difference between the five-point average value and the current value is higher than a new set value (comparison threshold) then a clap is detected, as described in the following equation. This algorithm yielded good and reliable results and was chosen.

$$a_{sqrt}[n] - \frac{1}{5}\sum_{k=1}^{5} a_{sqrt}[n-k] > threshold \tag{3}$$

where $a_{sqrt}[n]$ is the current acceleration sample and $a_{sqrt}[n-k]$ represents the k previos sample,

6) Pure Data on Raspberry Pi: The open-source visual programming environment, Pure Data version 0.49-1 was chosen due to availability (open-source)[8], philosophy from their creators

> "... to put the power of computing [...] into the hands of people all over the world"[9]

and due to its capability of handling real time audio, send/receive OSC messages, store experiment data, and the option to display visual feedback via HDMI. Pd is compatible with the native hardware and software of the microcomputer RPi, which is capable of running Pd patches in real time. It was decided to proceed with a Raspberry Pi 3 model B which comes equipped with RJ45 Ethernet, WiFi, Bluetooth,

[8] http://www.puredata.info
[9] http://www.raspberrypi.org

digital-to-analog audio conversion mini jack output, and full-HD HDMI output (to show game scores and feedback). The RPi can run headless (without an attached display); in this case all results will be presented via the ESP running the web server (future work).

The Pd patch comprising the game and analysis software incorporates the native OSC libraries with the mrpreach-vanilla external. It contains the audio engine playing sounds for the user, it handles the OSC message scheme, and performs data recording and saves the data via our storage system. It also translates the conditional game logic received from the VM via OSC and computes an overall game score for the players (see figure 9).

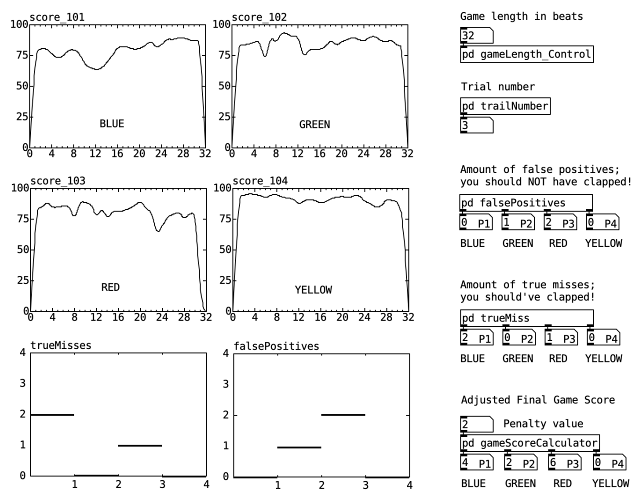


Fig. 9: Excerpt of the Pure Data patch showing game scores and penalties.

Initially, the Pd patch was developed to run on a Raspberry Pi 1. The patch is fully functional on the Raspberry Pi 1, but upgrading to a newer version seemed more beneficial due to CPU architecture limitations (not processing power). The Raspberry Pi 1 comes with an ARM1176JZ(F)-S CPU which architecture is not supported by Node.js anymore. For future work, we would like to send and receive OSC messages via a web server hosted on the RPi. Node.js is able to handle OSC messages via HTTP. For future work the web server could be hosted on the RPi running an Apache server with PHP, Node.js, Socket.io, and automata's osc-web[10].

III. Experiment design & method

A. Setup

To explore the link between gamification in a multiplayer setting and rhythm training, an experiment has to be conducted. The main target of this experiment is non-musically trained participants, however, we did not discriminate participants based on their musical ability. 20 participants (19 men and one woman, aged between 21 and 47) in five groups of four were recruited throughout the AAU campus. The experiment takes approximately 25 minutes. Participants are compensated with food and beverages. They were also asked if they would agree to being filmed and photographed during the experiment with no objections.

This experiment was divided into three stages: pre-intervention baseline test, game (intervention) and post-intervention test. Before and after the experiment, the participants are to fill out two questionnaires. A group of four participants is brought to an isolated room where they are instructed how to operate the wearable devices. Before carrying out the pre-intervention baseline test, they are asked if the wearable device feels comfortable and are encouraged to move the device around to ensure it is non obtrusive. Ensuring comfortability for every participant, before moving on to the actual experiment helps mitigate performance issues that may arise. In the case where a full group of four participants cannot be recruited, the researchers will join the group to fill up the vacant spot(s), which allows for a more interactive player experience. Fortunately, from the obtained data samples, no groups required this type of researcher intervention.

Furthermore, the participants are made aware of the scoring system. Each wearable device's LEDs light up whenever a clap is detected, with scores ranging from 0 to 100 being paralleled in the LEDs lighting up from red to green (i.e. the greener it is the more positive the result). The flow of the game is presumed to improve by providing the participants with instant feedback on their performance. Figure 10 illustrates the scoring system described.

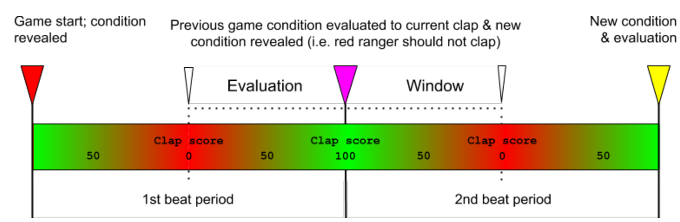


Fig. 10: Scoring system for rhythm game. Red arrow indicates game condition 'Red player do not clap on the next beat' with the LEDs on the VM flashing red on the first quarter-note. The following quarter-note, a pink arrow indicates when the red player's 'do not clap' condition gets evaluated while also showing the next game condition; pink colour indicating that everyone should clap on the following quarter-note.

The participants are instructed to attempt to clap on beat as close as possible to a series of repeating audio loops which will be played during the two tests and the rhythm game. These audio files were time stretched with Ableton Live Suite 10.1.5 utilising the Beats Mode time-stretch algorithm and are presented in 44.1kHz 16Bit normalised WAV format through a wired JBL Charge 3 loudspeaker. The original tempo of the audio loops is 90 BPM.

B. Baseline test; pre-intervention

Prior to beginning the initial baseline test, the participants are requested to fill out the first questionnaire[11] in regards to their personal details (age, sex), for statistical purposes only; their experience with music and rhythm related activities (e.g.

[10] git://github.com/automata/osc-web.git

[11] https://forms.gle/J1WqzhPM7sQmqV5V9

if they play an instrument, do they play rhythm games, if they dance, etc.) and music genre preferences. This questionnaire takes inspiration from the Goldsmith's Musical Sophistication Index (Gold-MSI [16]) to help determine one's musical and rhythmic abilities. The participants are also asked if they know one another as it could be an interesting factor to explore if interpersonal connections have effects on their performance. However, from the acquired sample data most participants knew at least one other participant in the group, not allowing for a proper evaluation of this parameter.

One assumption is that participants with more musical experience and behaviours will perform better since it can be assumed that these participants might have better rhythmic perception. Some correlation could possibly be found between a participant's favouring towards more beat driven musical genres such as house and techno music or hip-hop and their performance during their game time. Each participant takes the baseline test individually in player number order. The test requires each participant to clap on tempo to two audio loops for eight bars (32 beats) at 90 BPM in a 4/4 meter: a basic metronome loop and a simple drum loop with little syncopation. The order of stimuli was randomised for each group. The participants clap only to the sound of these audio loops, without a visual metronome. Utilising two distinct sounds can help the participants familiarise themselves with the game as well as provide some insight on how a loop with subdivisions can influence the results compared to the ones obtained from the quarter-note metronome. Having both a simple metronome and a drum loop was inspired by the Beat Alignment Test which is a tool to test "beat processing abilities in the general population" [8]. There is a 4 beat vocal count-in (one, two, three, four) before starting each trial. The remaining participants stay outside the room with some of the researchers until called for their turn in the baseline test.

C. Rhythm game

Upon completion of the first test, all participants are called into the room to begin the multiplayer rhythm game. Once again, the participants are asked if there have been any comfortability issues and if they are ready to proceed. If any issues are present, the researchers should try to aid in resolving these, then the rules of the game are explained. Eight rounds of eight bars of four distinct sounds (metronome and simple drum loop from baseline test as well as a more syncopated loop and finally uptempo breakbeat with many syncopated subdivisions, see figure 1) all in 4/4 meter are played, firstly all four in order at 90 BPM and then at 100 BPM. The VM is introduced and visible to all participants, flashing its LEDs in six distinct colours, each with a different purpose: when the lights flash pink, every participant claps on beat; when white flashes, all participants should skip the following beat (i.e. they do not clap on the beat following the white flash) and when the VM flashes one of the participant's respective colours (red, blue, green or yellow) that participant should skip the following beat (i.e. if red flashes, the participant representing the red colour should not clap on the next beat) while the remaining participants clap as normal. The participants are introduced to the game by running a test round to allow them completely grasp the game rules. Figure 11 provides an explanation of the game conditions.

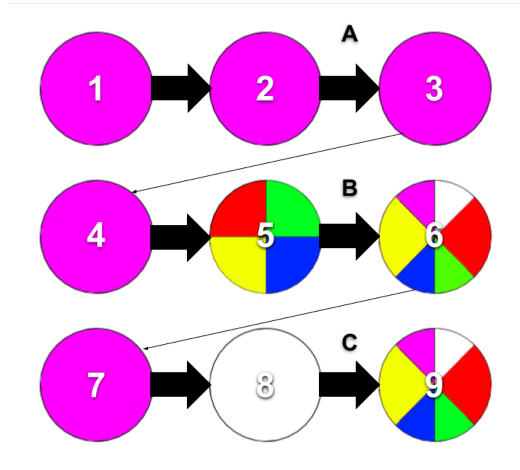


Fig. 11: The conditions for the rhythm game; colours represented are displayed in the VM: condition A displayed on beat 2 = every player claps on beat 3; condition B displayed on beat 5 = the player that had their colour flash on beat 5 does not clap on beat 6, regardless of the colour that shows up on beat 6, every other player claps; condition C displayed on beat 8 = no player claps on beat 9, regardless of the displayed colour on beat 9.

Once the instructions are presented, the researchers are to ensure the participants understood the rules to the best of their abilities before proceeding. It should be stated that two more conditions are present for the scoring system: when a participant misses a clap (true miss) or when a participant claps when not supposed to (false positive), two points are deducted from their final score every time these conditions apply. The participant's drift scores adjusted to account for any penalties are presented at the end of every round on the dedicated web server for every participant to see (see figure 9.

D. Baseline test; post-intervention

After the game, the baseline test is redone by each participant. Before taking the test, the participants are asked to fill out the second questionnaire[12] regarding the game itself and their overall experience playing the game. The questions cover: if the participants had any issue during gameplay making the experience less enjoyable or if there was any especially difficult part during the game, 5-point scale questions (from strongly disagree to strongly agree) regarding the comfortability of the device, the fun and difficulty factors of the game, and if instructions were properly explained. Participants were also asked if they used any strategies to help them keep tempo (e.g. bobbing their head to the rhythm, tapping their feet, dancing). The final part of the questionnaire regards their confidence in their musical abilities: if they were perceived to have played a role in the final scores (5-point scale), if their confidence improved upon playing the game and if so, how likely they

[12] https://forms.gle/DfsjoNkpJwbB3mo46

are to engage in musical activities in the future and how much they enjoyed the game, both in a 10-point scale. There is also an open feedback section.

IV. Data Analysis & Results

After running the experiment, the recorded data can be analysed. We looked at local and drift variability, improvements between both pre- and post-intervention tests and data collected from both questionnaires. We look for improvement by comparing drift scores from pre- to post-intervention tests and cross-referencing them with the answers to the perceived difficulty of the game. Upon performing linear regression on the data, the following trends were observed:

- Participants with least or negative improvement were the ones who found the game most challenging;
- Participants with higher game scores were also found to have the least improvement (pre- and post-intervention drift score difference close to zero);
- Participants who claimed to be less experienced with rhythm games and have fewer interactions with rhythmic performance showed the most improvement.

When analysing the difference in average local scores between both tests for the metronome sound, we found a decline of 0.14%. In contrast, the difference in average local score for the simple drum loop showed an improvement of 5.83%. An interesting result is that the scores were higher and less fluctuating for the simple beat drum loop than the metronome sound. The difference in averages between the simple beat drum loop and the metronome sound was 17 % (12 point difference out of the 100 point scale) in the pre-intervention test and 24 % (17 points) for the post-intervention test. The standard deviation (σ), comparable to the consistency of even claps between all participants, a score related to the overall drift, was lower for the simple beat drum loop than the metronome for both tests. See table I for the results.

TABLE I: Average score and standard deviation for pre & post-intervention tests with different sound samples.

	Pre-intervention		Post-intervention	
	Metronome	Simple beat	Metronome	Simple beat
Average	69.77	81.99	69.67	87.07
σ	4.98	4.02	5.26	3.99

A decrease in average local score is also observed when comparing the overall average of each participants local clap scores for the full 32 beats of both the pre-intervention and post-intervention test. Average of all participants local clap score. For the simple drum loop, higher average scores was observed for both the pre- and post-intervention test. When comparing average local scores between the pre- and post-intervention test for the simple drum loop, a larger difference was observed than when comparing with the metronome sound. All average drift scores for the tests showed a declining trend line, as seen in table II.

Participants show an improvement in performance (difference between post-intervention and pre-intervention test) for the simple beat drum loop opposed to the metronome sound, which can be seen in figure 12. A difference of around 10% was found between the two sounds. An interesting tendency was found where player performance decreased over time as seen in figure 12. This tendency was consistent throughout all trials between both tests, however this decline becomes more pronounced in the post-intervention trials, especially for the metronome sound.

TABLE II: Trend line equations for pre- and post-intervention tests.

	Metronome	Simple beat
Pre-intervention	y = -0.414x+77.2	y = -3.03x+87
Post-intervention	y = -0.574x + 79.9	y = -0.315x+92.2

It is also worth mentioning that the vast majority of first claps of the participants just after the countdown were very accurate. The average value for the first clap was usually close to a 90 point score. Since the stimuli were presented in randomised order for each group, we assume this difference in results between sounds is a matter of the sound itself and not the order of presentation. Error bars of ±1% are shown on the graphs, since the score values are captured as decimals and then rounded to the nearest integer.

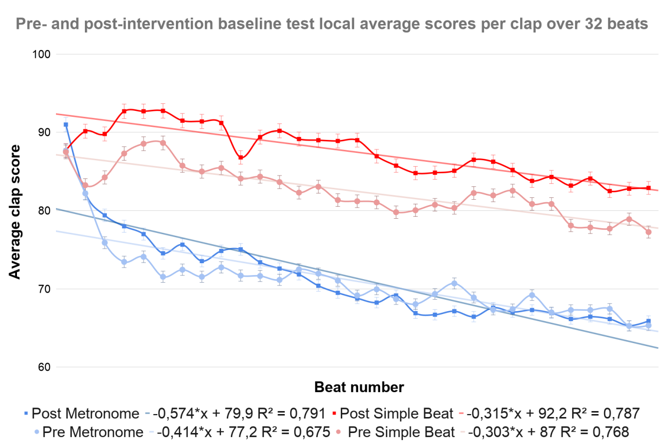


Fig. 12: Tendency over time of local scores decreasing for all trials in both baseline tests.

Comparing the players' average performance between the trials at 90 and 100 BPM (figure 13), it can be seen that in general players perform better at the higher tempo. This reason could be due to entrainment from the previous rounds as players would get more accustomed to the game's conditions during this later half of the game. We still see the same tendency as before, that players get a much lower score when following the tempo when hearing the metronome. When the game was played with the drum audio loops the score increased approximately by 20 points on average from the scores attained in the metronome rounds.

For both tempi, the performance in the first beats for the metronome trials was quite good (score around 90) but after about 4 beats, the scores had a tendency to decline to around 65 points. One subject group performed relatively well during the metronome rounds, however these results are

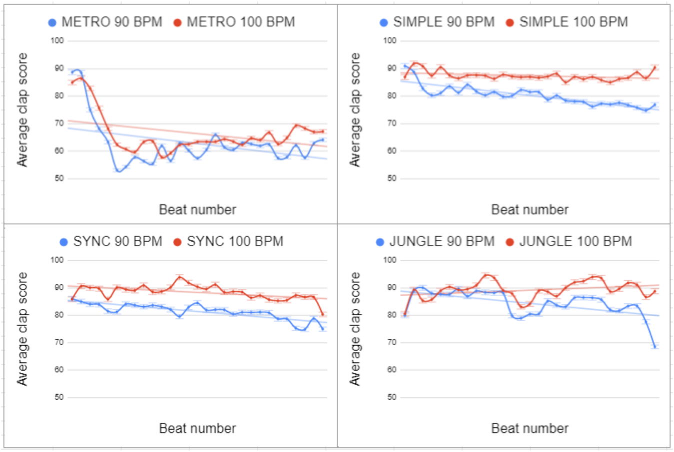


Fig. 13: Comparing player performance between game trials done at 90 (Blue) and 100 (Red) BPM. From top-left to bottom-right: metronome, simple drum loop, syncopated drum loop and jungle drum loop.

overshadowed by their overall performance as this was the highest scoring group (overall average drift score within the group = 84.16).

Figure 14 and table III display a trend over the audio loops at both game tempi. It can be seen that player performance seems to be more dispersed on the metronome loops, having no subdivisions. The simple and syncopated beats seem to have the most consistency (lower values for standard deviation) in performance with these loops having more subdivisions. For the jungle beat loop the consistency increases again, as it is more difficult to be accurate in tempo as there are many subdivisions in the sound. It can be concluded that loops with too many or too little subdivisions are harder to keep synchronisation.

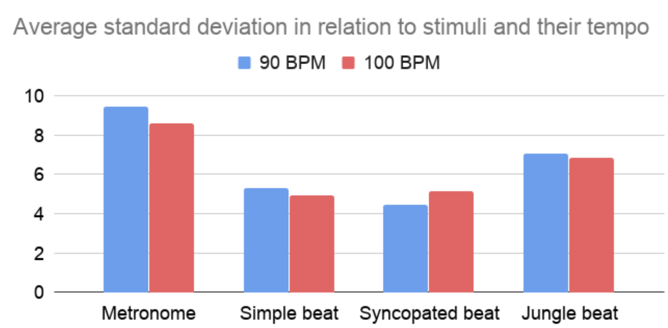


Fig. 14: Average of the standard deviations for all the players during a game with different tempos and sound samples.

TABLE III: Average standard deviation of all players for each sound and tempo.

	Type of beat			
	Metronome	Simple	Syncopated	Jungle
90 BPM	9.41	5.34	4.44	7.06
100 BPM	8.59	4.91	5.15	6.85

A. *Participant feedback*

From a total sample size of 20 participants, it was found that:

- All except one participant found the provided game instructions were clearly presented;
- 80% of the participants found the game had some challenge to it;
- 17 participants found the game device was comfortable to wear;
- The majority of participants did some action to aid them in keeping synchronisation, with preference to foot tapping (11 players) and bobbing their head (9 players).

Although not as high a number as hoped, six participants agreed that their confidence in their musical skills had improved upon partaking in this experiment, and nine participants felt very inclined to participate in more musical activities in the future. Overall, some participants claimed they faced some difficulty when the rhythm game introduced the VM due to the complexity of the rules, however every participant claimed they had fun playing the game and enjoyed themselves throughout.

V. Conclusion & Perspectives

In developing a wearable and wireless game device capable of capturing high resolution real-time performance data to be played in a competitive multiplayer game while doubling as a psychological experiment, we consider the project a success. Local variability in the baseline tests for the simple drum loop showed an average standard deviation across all clap scores which decreased slightly from the pre- to post-intervention test. The overall average score increased from 81.99 points to 87.07 points. The same comparison for the metronome sound showed an increase in overall average standard deviation (4.98 for pre-intervention test and 5.26 for the post-intervention test). The overall average score was almost identical for pre- and post-intervention test (69.77 for pre-intervention and 69.67 for post-intervention). For the game at 90 BPM, a decrease in standard deviation was observed for the first three trials: 1. metronome (σ = 9.41) $\rightarrow$2. simple beat (σ = 5.34) $\rightarrow$3. syncopated beat (σ = 4.44), followed by the fourth audio loop, the breakbeat with many syncopations (σ = 7.06). The breakbeat which was deliberately chosen as the hardest level game trials. At 100 BPM, a somewhat similar tendency was observed: decrease from metronome (σ = 8.59) to simple beat (σ = 4.91), then slight increase from simple beat to syncopated beat (σ = 5.15), and a higher increase for the breakbeat (σ = 6.85) (see table III for an overview).

The results of the experiment hint at various interesting correlations, but given the small sample size of participants, and how and where they were recruited, a more representative experiment with a larger and more diverse sample size would be preferred. In order to verify the effect of rhythm training more thoroughly, a control group whom did the baseline test twice within a couple of days without playing the game would have been needed. More trials and training over a longer period, say a couple of weeks, would likely show stronger overall results. It was complicated to do correlations between some of the qualitative data from the questionnaires (e.g. musical skills, behaviours, and musical genre preference) to the quantitative test and game data, but some interesting correlations were found; higher pre- and post-intervention test scores correlated with higher game scores showing less improvement. The opposite was found as well.

Training in a competitive style with pseudo collaborative tasks supported group cohesiveness according to participants. Interestingly when considering the overall task of the game was competitive. This supports the choice of the game condition, that universal task relations (all participants do not clap) increase perceived group unity.

Six out of 20 participants responded positively to an increase in musical abilities and nine out of 20 would like to engage more in music activities in the future.

Did the players enjoy themselves, absolutely. Is there room for improvement, of course.

A. Future work

A future version of the game / test device could incorporate more game modes, more levels, more sounds, advanced features like sound and timbre control, a visual metronome incorporated in a 2D platform game hosted on the RPi showed on a connected HDMI monitor. The device could also double as an OSC / MIDI compatible remote control which you could link to a Digital Audio Workstation (DAW) for expressive music performance and production (like the MI.MU Glove). If the audio engine gets ported to Faust, attaching a speaker to the device would make it usable for individual or group training for patients with neurodegenerative motor diseases.

Acknowledgments

The realisation of this project would not have been possible without the help of our supervisors Sofia Dahl and Dan Overholt and their extensive insight, the AAU E-lab staff members Jesper Greve and Peter Williams for their continuous technical assistance and infinite supply of machinery and Razvan Paisa for his ESP32 expertise, love and support.

References

[1] V. Bégel, A. Seilles, and S. Dalla Bella, “Rhythm Workers: A music-based serious game for training rhythm skills,” *Music & Science*, vol. 1, Jan. 2018.

[2] J. Lee and J. Hammer, “Gamification in Education: What, How, Why Bother?,” *Academic Exchange Quarterly*, vol. 15, pp. 1–5, Jan. 2011.

[3] V. Bégel, I. Di Loreto, A. Seilles, and S. Dalla Bella, “Music Games: Potential Application and Considerations for Rhythmic Training,” *Front. Hum. Neurosci.*, vol. 11, p. 273, May 2017.

[4] N. P. M. Todd and C. S. Lee, “The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives,” *Front. Hum. Neurosci.*, vol. 9, Aug. 2015.

[5] B. H. Repp, “Sensorimotor synchronization: A review of the tapping literature,” *Psychonomic Bulletin & Review*, vol. 12, pp. 969–992, Dec. 2005.

[6] N. Ehrlé and S. Samson, “Auditory discrimination of anisochrony: Influence of the tempo and musical backgrounds of listeners,” *Brain and Cognition*, vol. 58, pp. 133–147, June 2005.

[7] G. Madison, O. Karampela, F. Ullén, and L. Holm, “Effects of practice on variability in an isochronous serial interval production task: Asymptotical levels of tapping variability after training are similar to those of musicians,” *Acta Psychologica*, vol. 143, pp. 119–128, May 2013.

[8] J. R. Iversen and A. D. Patel, “The Beat Alignment Test (BAT): Surveying Beat Processing Abilities in the General Population,” p. 5, 2015.

[9] D. J. Getty, “Discrimination of short temporal intervals: A comparison of two models,” *Perception & Psychophysics*, vol. 18, pp. 1–8, Jan. 1975.

[10] A. Friberg and J. Sundberg, “Time discrimination in a monotonic, isochronous sequence,” *The Journal of the Acoustical Society of America*, vol. 98, pp. 2524–2531, Nov. 1995.

[11] L. Holm, F. Ullén, and G. Madison, “Motor and executive control in repetitive timing of brief intervals.,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 39, no. 2, pp. 365–380, 2013.

[12] M. Csikszentmihalyi, *Finding Flow - The Psychology of Engagement*. BasicBooks, 1997.

[13] M. Cziksentmihalyi, “Flow – The Psychology of optimal experience,” p. 8.

[14] J.-S. Lee, Y.-W. Su, and C.-C. Shen, “A Comparative Study of Wireless Protocols: Bluetooth, UWB, ZigBee, and Wi-Fi,” in *IECON 2007 - 33rd Annual Conference of the IEEE Industrial Electronics Society*, pp. 46–51, Nov. 2007. ISSN: 1553-572X.

[15] C. Arnould, “Clap detection with Microcontroller,” p. 22.

[16] D. Müllensiefen, B. Gingras, L. Stewart, and J. Ji, “Goldsmiths Musical Sophistication Index (Gold-MSI) v1.0: Technical Report and Documentation Revision 0.3,” p. 69.