Designing Games with a Purpose for Data Collection in Music Research. Emotify and Hooked: Two Case Studies

Anna Aljanaki[1], Dimitrios Bountouridis[1], John Ashley Burgoyne[2], Jan Van Balen[1],
Frans Wiering[1], Henkjan Honing[2], Remco Veltkamp[1]

[1]Utrecht University, Princetonplein 5, 3508 TB Utrecht, Netherlands
{A.Aljanaki, A.Bountouridis, J.M.H.VanBalen, F.Wiering,
R.C.Veltkamp}@uu.nl
[2] University of Amsterdam, Science Park 105, 1098 XG Amsterdam, Netherlands
{J.A.Burgoyne, Honing}@uva.nl

Abstract. Collecting ground truth data for music research requires large amounts of time and money. To avoid these costs, researchers are now trying to collect information through online multiplayer games with the underlying purpose of collecting scientific data. In this paper we present two case studies of such games created for data collection in music information retrieval (MIR): Emotify, for emotional annotation of music, and Hooked, for studying musical catchiness. In addition to the basic requirement of scientific validity, both applications address essential development and design issues, for example, acquiring licensed music or employing popular social frameworks. As such, we hope that they may serve as blueprints for the development of future serious games, not only for music but also for other humanistic domains. The pilot launch of these two games showed that their models are capable of engaging participants and supporting large-scale empirical research.

1   Introduction

Music Information Retrieval, or MIR, is a fast-developing interdisciplinary scientific field that treats music with data-mining techniques. These research methods often entail the collection of large ground-truth datasets. For some tasks, such as tonality or chord labeling, musical experts are needed to perform the annotation task. For other tasks, such experts are unnecessary and sometimes even undesirable. This is the case, for instance, when annotating music with tags or emotions, or measuring music similarity, because for these tasks researchers are more interested in variation in listeners' perception in general, as opposed to the theoretically more consistent opinions of experts. This second category of tasks is well suited for crowdsourcing data collection, for instance by using serious games.

Serious games, i.e., games that have non-entertainment purposes, have found multiple applications in health-care [5], education [8], and professional training [16]. Serious games are often perceived as a type of edutainment [17], for which there can be a diversity of goals: acquiring new skills, theoretical knowledge application in a simulation of a real world situation, or even informing oneself about a particular polit-

ical situations. Serious gaming comprises all games that pursue goals other than entertainment [17]. In this paper, we will concentrate on non-educational type of serious games, which normally have the purpose of gathering data from participants as a form of crowdsourcing. These games are called 'games with a purpose' (or GWAPs). Such games are created in order to provide a framework where humans perform tasks that computers cannot (yet) solve alone. The most prominent example is, perhaps, the ESP game – a GWAP for image labeling [1]. The data collected by ESP game have been used to improve image search and image recognition algorithms. In the field of MIR there are many tasks where human-computer collaboration is necessary. In this paper we are presenting two GWAPs that were created to collect metadata for music: Emotify, for emotional annotation of music, and Hooked, for studying musical catchiness.

1.1 Motivation

There are several reasons why games with a purpose are especially suitable for data collection in the realm of music. Firstly, for many tasks involving music, the common-sense expertise that every adult music listener possesses is sufficient, regardless of whether these listeners have had formal training in music. Secondly, listening to music is pleasant and self-rewarding. Most people enjoy listening to music, and therefore it is easy to create engagement. The third reason is that sometimes it is simply infeasible to collect data in other ways. Musical tasks are usually very time-consuming and the responses are subjective, and thus music experiments require a lot of time to complete and large numbers participants to control for presence of the effect(s) in question. It is sometimes possible to hire people to do the tasks in a way that Mechanical Turk [15] does, or alternatively to involve volunteers like The Open Mind Initiative [19] did, but these strategies can be very expensive because of the time-consuming nature of the tasks in first case or face the difficulty of verifying the accuracy of the results in second case. Games with a purpose, in contrast, are designed in such a way that the winning strategy is to provide the most correct and precise result possible. It is not possible to exclude vandalism or errors entirely when dealing with human-provided data, but as designers of GWAPs, we tried to minimize the risks, and we will discuss the techniques that we employed for that below.

2 Related Work

The term and concept of 'gaming with a purpose' was first suggested by Luis von Ahn, a pioneer in the area of human-based computation games [1], who introduced the ESP game in 2004. The ESP game is a competitive two-player game, whereby people provide labels for the pictures and score points by guessing the same answer. Google purchased a license to create its own version of the game in 2006. In 2008, a similar game called TagATune was created to enable music annotation with tags [12]. TagATune was designed to produce tags that would be much less subjective than those one could obtain from social music websites like last.fm. In TagATune, a player is randomly paired with a partner, both of whom must label a short (thirty-second)

musical excerpt with a series of tags. Based on their opponent's tags, players must guess whether they and their opponent have listened to the same song or not. In such a setup, tags referring to personal musical taste or subjective associations with music will naturally be avoided by players, as such tags are unlikely to provide useful information to a random opponent. MajorMiner and HerdIt! [14, 2] are similar to TagATune in design and purpose, but HerdIt! uses Facebook as a platform and supports multiplayer games.

Apart from GWAPs that collect textual annotations, some GWAPs for music have also collected other metadata, such as emotional annotations (MoodSwings) [10]. In MoodSwings, players jointly listen to the same musical fragment and provide continuous annotations by pointing with a mouse at a certain location on the screen, where the screen presents a two-dimensional representation of an emotional model.

3 Emotify: Collecting Emotion Annotations

Many MIR research areas are stimulated by music industry demand, and automatic music emotion recognition is one of them. It is easy and natural for people to organize music by emotion: witness the popularity of *stereomood.com*, a website where people collaborate to create emotion-based playlists. There is a need for technology that could automate this process. Automatic music emotion recognition relies on ground truth data, however, and there are no sizeable public datasets that could be used for training of the algorithms. The tags that can be collected from social music tagging websites, for example, lack consistency.

Moreover, there are two distinct goals for automatic music emotion recognition that need to be handled differently. The first task is automatic selection of background music: music to accompany a film or a commercial, with a requirement to express a certain emotion. Secondly, a musical piece might be selected as a means of mood regulation, for instance in music therapy settings or as a background for physical exercises. In the first case, music expresses a certain emotion; in the second case, music induces an emotion in listeners. The second case – *induced emotion* – is in the scope of our article and the game in question.

3.1 Emotify: Design Decisions

When discussing related work, we mentioned MoodSwings, a game for emotional annotation of music. MoodSwings focuses on *perceived* emotion (as opposed to *induced* emotion), which is also apparent in their choice of emotional model and their method of data collection. In our game, in contrast, we are trying to collect *induced* emotion annotations. For a game with a purpose, induced emotion creates a design problem. A standard type of player engagement in GWAPs is making players compete over giving the most standard answer possible (which is also supposed to be the correct one). In case of *induced* emotion there is no correct answer, and it would be misleading to encourage the listener to look for one. Induced musical emotion is by its nature personal and subjective.

This is why we introduced a different fun element than competition. We decided to create engagement by providing a feedback on player's answers in the manner of a psychological quiz. In addition, we decided to use a social network in order to give the player a possibility to compare his musical tastes and perception to those of his or her friends in the social network. There were three feedback elements in the game.

- **Continuous** feedback: A score calculated as a correlation of the player's answers to the averaged answers of other players. This score is recalculated after every answer and averaged over all answers.
- **Final** feedback: A histogram of the player's emotional responses for songs that player liked and for songs that the player disliked. This feedback was provided only after completing 10 songs, and thus we stimulated user to continue by promising a reward.
- **Playlist** feedback: Feedback on every song to which the player listened. Players had the possibility to listen to the whole song (not just the initial one-minute excerpt), and to see a detailed comparison to other players.

We hope that by designing such a feedback scheme, we encouraged players to give sincere and serious answers and at the same time provide a reward for their contribution.

3.2 Game flow

We created two versions of the game: one as a Facebook application (http://apps.facebook.com/emotify/) and another as a stand-alone version (http://emotify.org/). Figure 1 shows a screenshot of the game interface. We chose to launch the game on the Facebook applications platform for several reasons. Firstly, it provides a possibility to gather background information (age group, gender, location) about the player without additional questionnaires. Secondly, it simplifies game dissemination. Thirdly, it gives the player a possibility to involve people known to the player, as it might be more interesting to be compared to a friend than it would be to people with whom one is not acquainted. Moreover, Facebook has already successfully hosted serious games, such as the Rapport Game for common-sense knowledge collection [11, 20].

The gameflow is as follows.

1. The player authenticates through Facebook (or alternatively, enters the game from the stand-alone website) and provides personal details: age, gender, musical preferences, first language, level of English, and current mood.
2. The player is randomly assigned to one of four musical genres (rock, pop, classical and electronic music) and can switch to any other if he or she so wishes. The player may also switch at any later time.
3. In every genre, the player is presented with a random sequence of musical excerpts, each one minute in length. If a player is invited by a friend through Facebook, he or she is presented with the same sequence as the player who sent the invitation. This constraint is necessary in order to enable comparison between them.

4. After listening to the one-minute fragment, the player selects up to three emotions from a list of nine. This limitation should encourage player to think more carefully about the choices he or she makes.
5. The player also may indicate whether he or she liked or disliked the music and whether he or she knows the song. The player may also provide a new emotion definition if none of the nine correspond to what he or she is feeling.
6. At any time, it is possible to skip listening and go to another song or another genre.
7. There is a countdown from 10 to 1, saying that after 10 fragments the player will receive final feedback on his or her emotional perception of music. The countdown should encourage players to listen to at least 10 fragments to earn a "reward". Players may continue after listening to 10 fragments, but we prefer them not to do so, because understanding emotional content of music requires concentration and sensitivity, which is difficult to maintain for long periods of time.
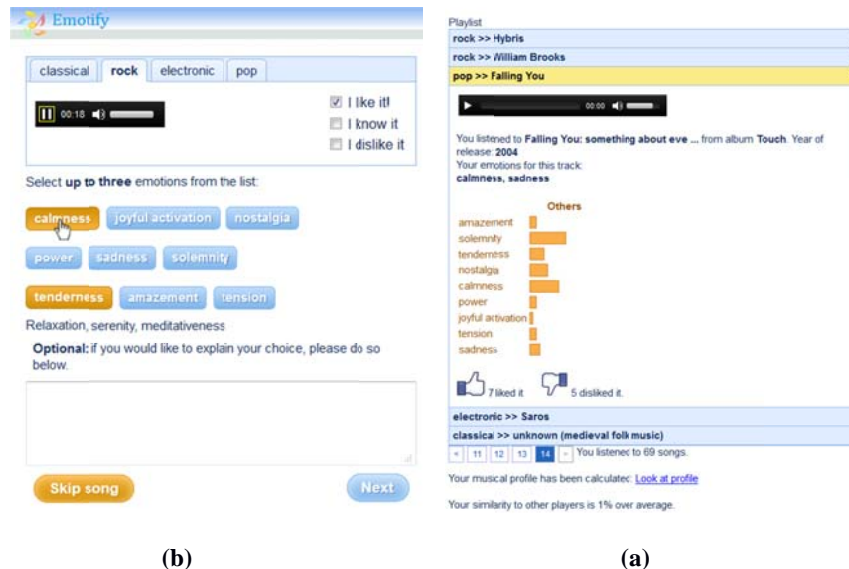

(b)   **(a)**

Fig. 1. Emotify screenshots: (a) game interface (b) feedback window

3.3 Discussion

Emotify was launched on the 1st of March 2013, and in the 4 months following the game launch, 1285 players played it. On average, they listened to 8 songs, and spent 13 minutes and 40 seconds playing the game ($SD = 12.62$). The actual time spent in the game differed a lot depending on the player. As we were advertising a game through online media, there were many players who merely examined the game and quit almost immediately, but there were also devoted players who spent a lot of time

listening to music. Overall, the players gave positive feedback to the game and were motivated by the reward scheme the game provided.

The data collected since game launch is openly accessible[1] will be regularly updated.

4   Hooked: A Serious Game for What Makes Music Catchy

Many listeners, even those without musical training, experience the pleasant feeling of recalling a song to memory after hearing a few seconds of its "hook". Similarly, listeners are immediately able to identify whether a new song will be "catchy". Understanding hooks and catchiness is valuable not only for studying musical memory, but also for MIR tasks such as recommendation and similarity. Nonetheless, little research has been conducted on these notions, perhaps due to the fuzziness of their definition and a lack of experimental data. From a cognitive point of view, we solve the first issue by defining a hook simply to be the most salient, easiest-to-recall fragment in a piece of music [9] and catchiness as the long-term musical salience. The second issue is addressed by Hooked, a game with a purpose aimed at studying musical catchiness. Our game is designed to support collecting data from the players, similar to serious games such as Foldit [7]. However, in contrast to the previous games, Hooked is developed for mobile devices, making use of their social nature for viral distribution and discovery.

4.1   Hooked: Design Decisions

Two scientific needs drove the design of Hooked. First, we needed to be able to work with well-known music, in order to capture fragments that truly remained in the participants' memories over the long term. Secondly, because each participant has his or her own listening history, we needed to be able to support a large number of participants for the sake of reliable statistics. Three tasks are central to how we transformed these scientific needs into entertaining gameplay paradigms: recognition, verification, and prediction. We will discuss them first outside the context of the game overall, and later as a unified game experience.

Recognition Task.
The recognition task plays the most important role in the game by triggering long-term memory. It is based on the following premise: the easier a music fragment to recall after a long period of time, the catchier it should be. Therefore, we devised a "game-based" – thus featuring "goals" – quiz-style game. Following the "drop-the-needle" paradigm, a piece of music starts playing from a random point in the middle while players are asked to recognize it within a fixed timeframe (e.g., $N = 10$ s). The theoretical literature on hooks suggests that hooks should coincide with the beginning of major structural sections (e.g., a new verse or chorus) [3], thus in our game, the

---

[1] http://www.projects.science.uu.nl/memotion/subset60/

starting points for each song are limited to manageable subset. Once the music starts playing, players are given two options ("Yes" and "No") corresponding to the prompted question "Do you know this song"? While the player listens to the music, points are counting down, penalizing players who listen longer than necessary. Based on that fact, penalizing players for "taking their time" actually motivates them to act as quickly as possible whenever their long-term memory is triggered by a hook. It should be noted that choosing "No" does not affect the player's score, since we do not want to encourage guessing. Figure 2 illustrates the recognition gameplay as implemented for iOS devices.
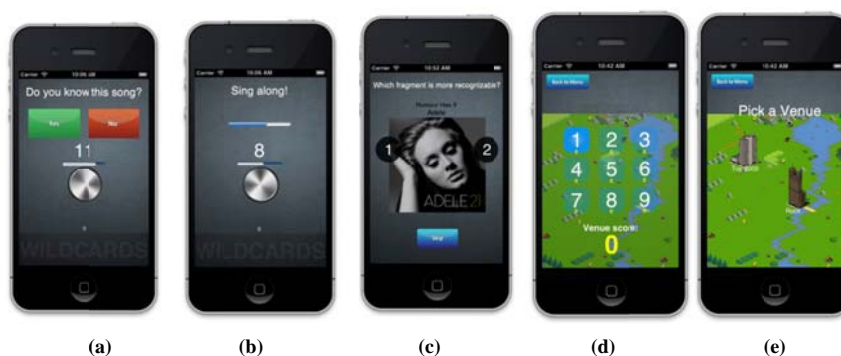

(a)	(b)	(c)	(d)	(e)

Fig. 2. Screenshots from the prototype. (a) The recognition task: A song starts from the beginning on an internal musical section, chosen at random, and the player must guess the song as quickly as possible. (b) The sound then mutes for a few seconds while the player is asked to follow along. When the sound comes back players must verify that the song is playing back from the correct place. (c) Occasionally, players instead must do the reverse: predict which of two sections is catchiest. (d) Ten recognition tasks constitute a level and groups of ten levels form a "Venue". (e) The player is asked to complete all "Venues" and their corresponding levels.

Verification Task.

The recognition task by itself is only valid for controlled environments. In addition, from a gameplay perspective, such a trivial task presents no interest for the player. Therefore, we needed a complementary task to verify the correctness of the player's answer. An initial implementation made use of the typical music trivia paradigm where players are asked to identify, or select from a number of options, the artist name, title or year of release of the song in hand. However, such an approach has two major drawbacks: it introduces context and it assumes that listeners always remember these facts about songs they know. Adding context would make guessing a viable option, which makes the game easier. Knowing a song on the other hand, and therefore recalling it musically, does not imply knowing its title or composer.

Our approach is based on the observation that once a music piece has been fully recalled to memory, players should be able to follow along even after playback

stopped. Therefore, as soon as the player hits "Yes" in the recognition task, the sound is muted while the song keeps playing for a fixed amount of time (e.g., $M = 3$ s). During this time, the player is asked to sing along (Figure 2b). When the sound returns there are two possible scenarios. Sometimes, the music will continue playing from the correct place. Other times, the playback can be offset by predefined number of seconds (e.g., $D = -15$ s). The player is then asked whether the music is playing from the correct place with the following question: "Is this what should be playing now"? A correct answer rewards the player with $N - response\_time$ points, whereas an incorrect answer subtracts $N - response\_time$ points from his or her current score.

Prediction Task.

Both of the aforementioned tasks were designed considering the notion of hook as a musical stimulus that eases recall. The prediction task, on the other hand, aims at capturing the listener's informal intuition of what is catchy or not. During this task ,the player is presented with two fragments (of size N) of the same song, accompanied by its title, artist and album cover (Figure 2c). The latter "metadata" are used to trigger the player's memory in case the musical stimuli fail. The player is asked to listen to both fragments and pick the one that he or she considers catchier. In the next part we describe how such a survey question can be meaningfully integrated inside the general game context.
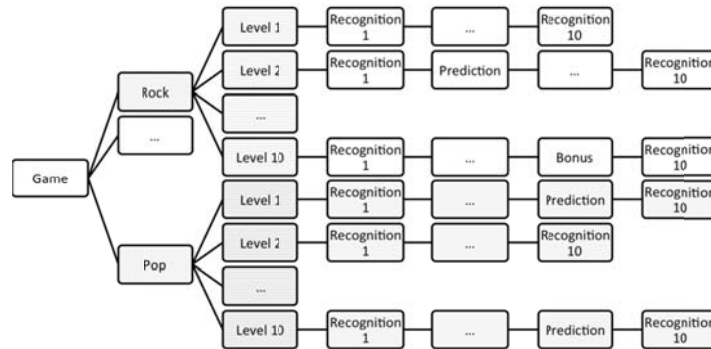

Fig. 3. The hierarchical layering of "Hooked!". The game is decomposed into "Venues", which themselves contain 10 levels each. Each level is completed after 10 recognition tasks, interrupted by prediction tasks and "bonus" rounds. Predictions appear with a chance of 20% while "bonus" rounds with a chance of 10%.

4.2 Hooked: Inside an Entertainment Game Context

The recognition-verification couple and the prediction tasks are independent tasks, i.e., no interaction between them occurs. Merging such data collection tasks inside a game structure was therefore challenging, especially considering that players are unaware of the underlying scientific questions they are helping to address. For them,

Hooked should be consistent with the current entertainment standards offered by popular mobile games. The next paragraphs describe our gameplay design approach and the incorporation of a social element.

Hierarchical Leveling.

One of currently popular mobile leveling schemes is based on the hierarchical arrangement of levels. At the top of the hierarchy, levels are grouped based on a common feature (e.g., scenery). These clusters, typically referred to as "episodes", commonly serve an underlying storyline progression. Episodes contain a fixed number of enumerated levels, which remain unavailable until all preceding levels have been successfully completed. Our game employs a similar structure. Instead of "episodes", however, we use the metaphor of "venues", which contain similar music in terms of either genre (e.g., pop or rock) or time (e.g., '80s or '90s). Each "venue" contains 10 levels, and each level consists of at least 10 recognition tasks. A level is complete if more than 6 recognition answers are correct. In order to increase difficulty as the player progresses, we prompt the player with older and less popular music at each new level. The corresponding features, release date and popularity (the so-called "hotness" feature) are available from the Echo Nest service [6].

Bonus Rounds.

As we previously mentioned, prediction tasks act as survey questions outside the game context. Incorporating them meaningfully was solved with "bonus" levels. Recognition tasks can be interrupted by prediction tasks. Each time a player completes a prediction task, the chosen fragment is stored in a special pool. The recognition task enters a bonus round periodically (unless the fragment pool is empty) for double points. The prompted fragment is randomly selected from the pool. Therefore, if the player picks the catchiest fragment in the prediction task, he or she will have higher chances to recall the song in the "bonus" level. Consequently, the player is motivated to perform the prediction task thoroughly. Figure 3 illustrates the hierarchical layering of Hooked, including predictions tasks and bonus rounds.

Wildcards.

Wildcards are a common scheme for increasing a depth of the gameplay and keeping players motivated. Our game rewards players with wildcards after "combo" recognition tasks (e.g., 5 correct recognitions in a row). Wildcards can be used during later recognition rounds for revealing the album cover of the prompted song, decreasing the mute time M, multiplying the gained points by a factor, and other strategies to help score more points.

Social elements.

Interaction between players, centered on the social nature of the mobile phone, has generated much interest among game designers in the past years. Due to the scientific purposes of our games, incorporating a social element aims at viral distribution and

discovery, thereby attracting a large number of players and statistically useful scientific data. Two frameworks were investigated: Facebook Games and Apple's Game Center. The first offers a convenient protocol for invites and social bragging on the news feed (e.g., "Check out my score on Hooked!"). The second offers a more direct interaction between players, such as turn-based and real-time matches, in addition to common leaderboards. As such, our current implementation employs Game Center to increase the game's depth and sophistication.

Hooked uses leaderboards and achievements as the primary step of social interaction. Players with a high score appear on the common leaderboard of the Game Center platform. Achievements are rewards for completing certain tasks and appear on player's Game Center profile. The most challenging interactions, however, are turn-based matches. Players can invite each other to a game of recognition tasks in a turn-based fashion. Our game allows challenges of up to two players. Each player must complete a recognition task before passing his turn to the other player. The prompted song for both players is the same, to ensure fairness. At the end of ten recognition tasks, the player with the higher score is announced to be the winner.

5   Hooked and Emotify: Musical Material and Game Settings

One of the most challenging tasks for serious musical games is acquiring a large set of licensed music data. Emotify and Hooked solve this problem in different ways. For Hooked, we needed music that would be widely recognizable for the majority of players; for Emotify, well-known music was, to the contrary, undesirable.

In order to avoid licensing problems, Hooked uses Spotify's iOS library, which offers a catalogue of around 20 million tracks [18]. In addition to its increased popularity and high audio streaming speeds, Spotify is partnered with Echo Nest such that the Echo Nest Analyzer can be applied conveniently to any item in the Spotify corpus. We use this link to obtain estimates of the start times of the major structural sections, year of release, and "hotness" for each song. Therefore, the Spotify–Echo Nest coupling presents a unique combination of vast and well-documented music data, proving to be an excellent choice for the back end of serious musical games.

As the main component serving our scientific purposes, the recognition task should be parameterized to separate catchy sections from the rest. Therefore the choice of mute time M, maximum recognition time N and offset D were of great importance. A pilot version, comprising of 32 songs and 20 participants, was run during the months May-June 2013 for that purpose. The configuration $M = 3$, $N = 10$ and $D = -15$ yielded the most statistically significant recognition time differences between sections [4].

In case of Emotify, the main purpose of the game is to collect induced emotion annotations. For musical perception, familiarity of the music has certain consequences, as associations with the music already known to listener might create an undesirable additional emotional response. This is why we decided to make use of relatively little-known music by the recording label Magnatune [13]. This label releases music under

Creative Commons licenses, which permits free use of its music for non-commercial purposes.

6   Discussion

In this article, we presented two musical games with a purpose: Emotify and Hooked. While Emotify has already been launched, Hooked is still in the pilot stage. Discussing the data collected by the games and scientific value of this data is out of the scope of this article; for more information on these questions, we refer readers to [4]. The games that we presented are serious games created for the purpose of data collection, or GWAPs, but in contrast with other GWAPs, they do not force players to compete to guess the correct answer. Emotify rewards players by giving feedback on their own input and comparing their answers to other players' input; Hooked stimulates players by awarding points for the correct answers. Hooked collects data in an indirect way by measuring players reaction time, whereas Emotify's data come from direct self-assessment. Furthermore, Emotify and Hooked use two distinct musical corpora and two different distribution channels: web-browsers and mobile devices.

Emotify uses social networks as a platform. We faced some limitations, however, when using the Facebook application platform. The Emotify game tried to use invitations to involve new players into the game, but the acceptance rate of the invitations was very low: of the invitations sent by players of Emotify, only 6% were accepted. Moreover, people were reluctant to use the Facebook version of the game. After running the pilot of the game on Facebook, we launched an independent website for hosting the game, and both the Facebook version and the independent website were advertised together. Having a choice, more than 90% of the players preferred to use the independent website. Hooked, on the other hand, uses mobile devices, a familiar platform with its own interaction patterns. As a consequence, users sometimes attempted to employ typical tactile gestures (e.g., pans or swipes) on the game elements. This resulted in some initial confusion, but with a few minutes of experience, the overall gameplay impression was clearly positive. All the participants in the pilot found the game rather addictive and fun. It should be noted that most participants showed rapid learning rate, meaning that their responses times were gradually decreasing as the game became more familiar to them.

7   Conclusion

We presented two case studies of serious musical games with a purpose. Data gathering inside a gameplay context has been addressed by two separate paradigms, each one employing a unique approach for transforming scientific needs into valid game practices. Two social interaction models have also been investigated: an invitation-based Facebook model and a competitive, turn-based match model. Preliminary investigations showed that data collected by the means of these GWAPs is of sufficiently high quality to support academic research [4].

Emotify and Hooked offer valuable insights regarding musical games and help to answer questions about music material acquisition, approaching player-participants, viral distribution, and more. In particular, they illustrate the importance of thinking outside the social-media box – an enormous proportion of players preferred *not* to integrate their serious gaming with social media – and the importance of including a modular and regular reward structure even for tasks like music that players normally enjoy. The approach described in this paper could also be used to build non-musical serious games.

References

1. von Ahn L., Dabbish L.: Labeling images with a computer game. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA (2004) 319–326
2. Barrington L., O'Malley D., Turnbull D., Lanckriet G.: User-centered design of a social game to tag music. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, New York, NY, USA (2009) 7–10.
3. Burns G.: A Typology of 'Hooks' in Popular Records. Popular Music, Vol. 6, No. 1 (1987) 1–20
4. Burgoyne J.A., Bountouridis D., Van Balen J., Honing H.: A game for discovering what makes music catchy. In *Proceedings of the 14th International Conference on Music Information Retrieval*, Curitiba, Brazil (2013)
5. Deponti D., Maggiorini D., Palazzi C. E.: DroidGlove: An android-based application for wrist rehabilitation. In *International Conference on Ultra Modern Telecommunications Workshops* (2009) 1–7
6. EchoNest. http://echonest.com/
7. Foldit. http://fold.it/portal/
8. Gaggi O., Galiazzo G., Palazzi C., Facoetti A., Franceschini S.: A Serious Game for Predicting the Risk of Developmental Dyslexia in Pre-Readers Children. In *21st International Conference on Computer Communications and Networks (ICCCN)* (2012) 1–5
9. Honing H. J.: Lure(d) into Listening: The Potential of Cognition-Based Music Information Retrieval. Empirical Musicology Review, Vol. 5, No. 4. (2010) 121–26
10. Kim Y., Schmidt E., Emelle L.: MoodSwings: A Collaborative Game for Music Mood Label Collection. In *Proceedings of the 9th International Conference on Music Information Retrieval*, Philadelphia, PA, USA (2008) 231-236
11. Kuo Y., Lee J.C., Chiang K., Wang R., Shen E., Chan C., Hsu J.: Community-based Game Design: Experiments on Social Games for Commonsense Data Collection. In *Proceedings of the ACM SIGKDD Workshop on Human Computation*, Paris, France (2009) 15-25
12. Law E. L. M., von Ahn L., Dannenberg R. B., Crawford M.: TagATune: A game for music and sound annotation. In *Proceedings of the 8th International Conference on Music Information Retreival* (2007) 361–364
13. Magnatune http://magnatune.com/
14. Mandel M., Ellis D.: A web-based game for collecting music metadata. Journal of New Music Research, Vol. 37, No. 2. (2008) 151–165
15. Mechanical Turk https://www.mturk.com/mturk/

16. Microsoft Flight Simulator. http://www.microsoft.com/games/fsinsider/
17. Ratan, R., Ritterfeld, U.: Classifying serious games. Serious games: Mechanisms and effects. Routledge (2009) 10 -24
18. Spotify http://spotify.com
19. Stork, D.G: The Open Mind Initiative. IEEE Intelligent Systems & Their Applications 14, 3 (1999) 19–20.
20. Takhtamysheva A., Krause M. Smeddinck J.: Serious Questionnaires in Playful Social Network Applications: In *Proceedings of Entertainment Computing*, Vancouver, Canada. (2011) 436–439