If You Wanna Be My Lover … A Hook Discovery Game to Uncover Individual Differences in Long-term Musical Memory

Korsmit, I.R.; Burgoyne, J.A.; Honing, H.

Published in:
Proceedings of the 25th Anniversary Conference of the European Society for the Cognitive Sciences of Music

If You Wanna Be My Lover… A Hook Discovery Game to Uncover Individual Differences in Long-term Musical Memory

Conference Paper · August 2017

3 authors:

Iza Korsmit
University of Amsterdam

John Ashley Burgoyne
University of Amsterdam

Henkjan Honing
University of Amsterdam

Proceedings of the 25th Anniversary Conference of the European Society for the Cognitive Sciences of Music, 31 July-4 August 2017, Ghent, Belgium
Van Dyck, E. (Editor)

If You Wanna Be My Lover… A Hook Discovery Game to Uncover Individual Differences in Long-term Musical Memory

Iza Ray Korsmit[1], John Ashley Burgoyne[2], Henkjan Honing[3]

Music Cognition Group, Institute for Logic, Language and Computation, University of Amsterdam, the Netherlands
[1]iza.korsmit@student.uva.nl, [2]j.a.burgoyne@uva.nl, [3]honing@uva.nl

ABSTRACT

Hooked on Music (Burgoyne, Bountouridis, van Balen, & Honing, 2013) is a citizen science project developed to uncover what makes music memorable. The game consists of two stages: firstly, *recognizing* a song fragment, secondly, *verifying* that one actually knows that song well, by correctly singing along to it. Half of the time, the music actually continues in the correct spot, half of the time it does not. This study aims to uncover individual differences in long-term musical memory and demonstrate the use of Multivariate Item Response Theory (MIRT; Hambleton, Swaminathan, & Rogers, 1991) for this purpose. We performed a MIRT-based exploratory factor analysis to research individual differences in the Hooked on Music data, and used demographic data and audio-based corpus analysis to interpret the results. To the best of our knowledge, MIRT-based analysis has not been used in this area before. For the recognition data, we found a four-factor model (R^2 = .47) that we interpreted as individual differences in music preference and/or attention to musical characteristics (*female vocals, rap/hip-hop, strong beat)*, and age (*music released in the 1990s*). Our verification models seemed to mostly be a result of the participants' direct task at hand, representing player's abilities to sing along with the music and catch big changes in the music fragments (correct continuation, R^2 = .09; incorrect continuation, R^2 = .14). This demonstrated the use of MIRT in uncovering individual differences and showed which factors are of relevance in long-term musical memory, although confirmatory analysis is necessary.

I. INTRODUCTION

The Spice Girls singing "if you wanna be my lover" will instantly trigger the memory of anyone who is somewhat familiar with the tune. They have been "hooked" into the popular 1990s hit "Wannabe" (whether they wannabe or not). Hooked on Music (Burgoyne, Balen, Bountouridis, & Honing, 2013) is a citizen science project developed to uncover what makes music this memorable. The data can help study which aspects of popular songs are more memorable than others, and due to its wide reach, they can also help study which factors influence *individual differences* in remembering these popular songs. Not everyone, of course, will become hooked on the Spice Girls' 1990s hit.

Hooked on Music is an online game that consists of two stages. First, in the *recognition* stage, the player hears a fragment from a pop song, randomly beginning at a structurally new part of the song (e.g., the beginning of a verse or chorus). The player has 15 seconds to decide if they recognize the song or not. If they recognize the song, the next stage comes into play: *verification*. When the player has answered "Yes" to the recognition question, the sound is muted for four seconds. When the sound is unmuted again, the player is asked whether the music continued in the correct spot or not. Half of the time the music does not continue in the right spot, i.e., if you were able to (correctly) sing along during the muted four seconds, your singing and the song are not in sync. This verifies whether the player really knew the song.

In previous research, Hooked on Music was able to show what the musical characteristics of a hook are (Van Balen, Burgoyne, Wiering, & Veltkamp, 2013). Our paper uses Hooked on Music's results to go a step further and uncover individual differences in recognizing popular music, using multivariate item response analysis (MIRT; Hambleton, Swaminathan, & Roger, 1991; Reckase, 1991). MIRT can be adapted to a form of factor analysis, where one takes into account the ability of the individual, as well as the difficulty of the item. In the case of Hooked on Music, this results in different factors that can explain individual differences in game performance, while acknowledging that not all song fragments themselves are equally easy to remember at baseline.

The literature points to several possible sources of individual differences. Our musical lexicon (Peretz & Coltheart, 2003) can be influenced by our *age* (reminiscence bump; Krumhansl & Zupnick, 2013, Schulkind, Hennis, & Rubin, 1999). This age effect was also apparent in the results of an older version of Hooked on Music, in which players were able to choose a playlist from one particular decade. Figure 1 shows that people who were born between, e.g., 1940 and 1949 were most likely to select a 1960s playlist; music from when they were 'in their reminiscence bump'. This trend continues in people born in later decades (people born before 1940 were very low in number, making those results unreliable)

Simply put, that to which you expose yourself will be that which you remember, so we can expect that individual differences in the game performance would also be influenced by a person's musical *preferences*. Besides division into (often arbitrary) musical genres, musical expression could also express itself as the five-factor model *MUSIC* (Rentfrow, Goldberg, & Levitin, 2011). A third possible source of influence might be the *level of processing* (Radvansky, 2011). That is, people have different listening styles, with some only treating it as background music and others immersing themselves in the music completely (Ter Bogt, Mulder, Raaijmakers, & Nic Gabhainn, 2010). This can influence recall and recollection (Craik & Lockhart, 1972; Hyde & Jenkins, 1973) and might influence *recognition* (similar to recall) and *verification* (which requires some recollection for the continuation of the song) differently. Finally, one might expect that music listeners would show individual differences in the *different aspects of music* they pay attention to. Not much research has been done in this area, but one study showed that, for example, people may be differently attuned to the beat (Grahn & McAuley, 2009; McAuley, Frater, Janke, & Miller, 2006). All these sources of individual differences

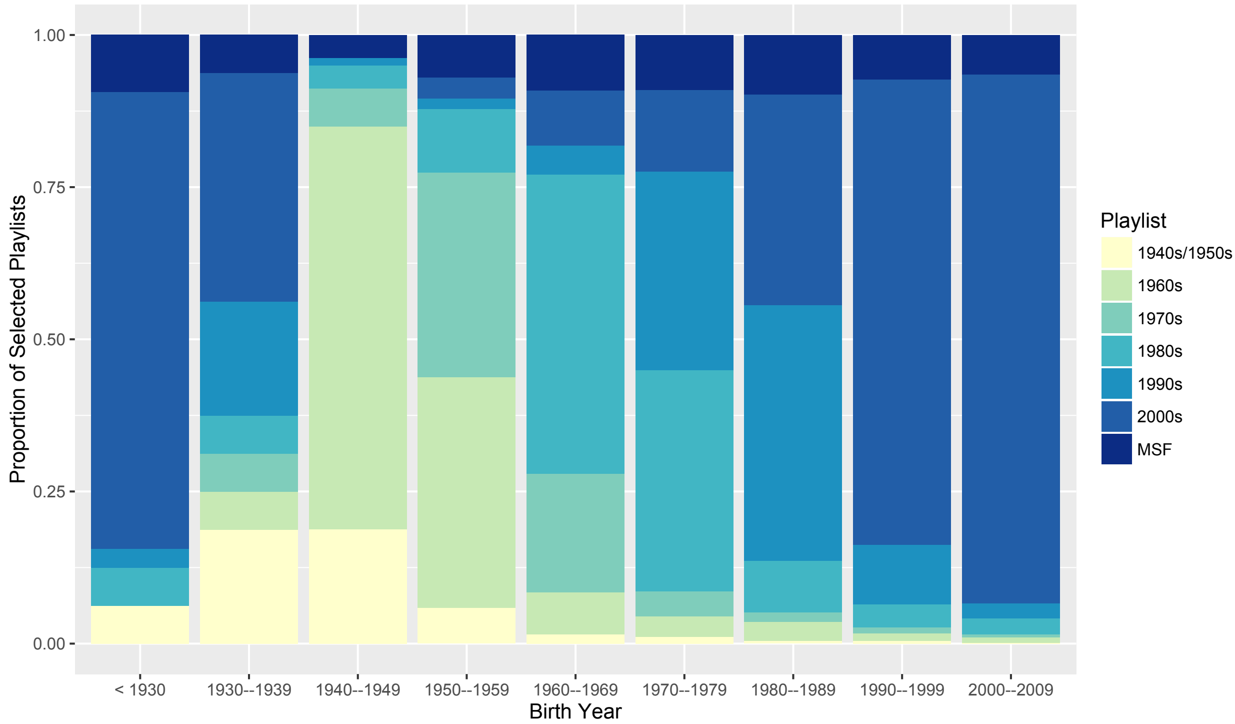

Figure 1. Proportion of playlists selected by Hooked on Music players

could also influence each other and be influenced by other individual differences like personality or socio-economic status.

The nature of this study is exploratory. We will use our literature findings as a framework for interpreting our results, but keeping in mind that many different factors could be at play. Our ultimate aim is two-fold: to uncover individual differences in long-term musical memory and to demonstrate the use of MIRT analysis in this type of research.

II. METHODS

A. Multivariate Item Response Theory

MIRT was developed in response to classical test theory. In classical test theory, it is not possible to separate a participant's ability from the difficulty of the test items. In classical factor analysis, these differences between test items would also be averaged out, whereas in an MIRT factor analysis, they are considered to be a source of information (Hambleton, Swaminathan, & Roger, 1991; Reckase, 1991). One can take into account not only the particular ability of a participant one is curious about, but also how a given test item might be easier to answer than another. Each song fragment of Hooked on Music is a test item, which will be more or less difficult to recognize or verify based on the ability of the player and the inherent difficulty of the fragment. This type of analysis has, to the best of our knowledge, not been used in long-term musical memory research before.

B. Factor Analysis

This study has an exploratory approach. Although we could make some predictions based on the literature reviewed above, we performed an exploratory factor analysis to uncover how many factors are at play in the recognition and verification stage. We based the number of factors to extract on several methods, as there is no one perfect method for doing so. We looked at parallel analysis, very simple structures (VSS), scree plots, and AIC_c values, to determine what number of factors was the most explanatory, without unnecessary complexity.

After factor extraction, we fitted the model to our output data. The factor loadings of the factors on each item, in this case, represent discrimination (a) parameters of the *item characteristic curve* (ICC; Reckase, 1991). These parameters represent the steepness of the item response curve at the point where the probability of a correct answer is .50. The steeper the curve – i.e., the higher the a-value – the more sharply an item differentiates between responders with high ability (θ) and low ability on that factor. In the case of recognitions, a high discrimination value for a song fragment on a particular factor indicates that the fragment distinguishes well between players scoring low and high on that factor with respect to whether the players will recognize the fragment; in the case of verifications, this is with respect to whether the players will correctly identify if the song continued in the right place. If the a-value is around zero, the test item does not distinguish well for a particular factor. With a negative discrimination value, the ICC is reversed: although an item also distinguishes well between people with low and high ability on that factor, a high ability increases the chances of *not* recognizing a test item or *not* being able to verify correctly.

Preliminary analysis showed that participants were very strongly inclined to answer that a song continued in the right spot, regardless of the correct answer. Therefore, correctly verifying that a song continued in the right spot proved to be much easier than correctly detecting that a song did not, indicating that these two tasks are not similar. This moved us to run the factor analysis three times: once for recognitions, once for verifications with *correct* continuations, and once for verifications with *incorrect* continuations.

C. Feature Analysis

To determine what these factors represent, we listened to the highly loading song fragments for each factor and determined what they had in common. To aid us in this listening process, we conducted two additional analyses. First, we correlated the demographic information with the predicted factor scores. Second, we extracted audio corpus description features from all the item-sets loading high on the factors with the CATCHY toolbox (Van Balen, 2016; Van Balen, Burgoyne, Bountouridis, Müllensiefen, & Veltkamp, 2015) and performed multiple regression analyses predicting the factor loadings.

D. Materials

All the music fragments were selected from the top 20 songs from the British pop charts of every decade since the 1940s. Songs for which a license could not be obtained were excluded from the selection. Each fragment started at a structurally new section of the song, as determined with the Echo Nest Analyzer (Burgoyne et al., 2013). For the verification stage, an incorrect continuation started 15 seconds after the time point of correct continuation.

E. Participants

From July 2014 to August 2015, 130.000 participants played Hooked on Music a total of 1.8 million times.[1] We could not take all these participants and plays into our analysis, because the quality of engagement varied widely. Data were excluded based on the following criteria:

- Any plays with recognition time under 1 s, which are not likely to be genuine attempts (0.2% of the plays discarded).
- Any participant that either always skipped recognition, always answered that a continuation was correct, or always answered that it was incorrect, which were also likely not genuine attempts (29% of the participants discarded).
- Any participant that played less than 50 song fragments. This enabled us to take only the most engaged participants' data into account, although there was no obvious cut-off point that distinguished engaged from casual participants (66% of the participants discarded).
- Any songs that were not appropriate for audio-based analysis, like Beatles Symphony Orchestra recordings that were used as proxies for unlicensable Beatles songs (7.2% of the songs discarded).

This resulted in a dataset of almost 140.000 plays (25% of the raw data), which came from 4082 participants and 184 song fragments. At the moment of analysis, the number of participants that answered all demographic questions was 494, but partially answered questionnaires were also used for analysis.

[1] hookedonmusic.org.uk

Most participants were born between 1990 and 1999 (n = 1264) or 2000 and 2009 (n = 513), and as few as 46 participants were born in 1960-1969. The gender distribution was also skewed, with 1569 females, 632 males, and 48 people stating "other" as their gender. Most participants had finished post-secondary education (n = 567) or master's education (n = 154). Most participants deemed their memory to be good (n = 846) or average (n = 675). Most participants also stated they listened to 1 to 3 hours of music daily (n = 984) and attended 1 or 2 music events annually (n = 873). Finally, most participants had either taken no music lessons at all (n = 671) or 1 to 5 years of music lessons (n = 703).

III. RESULTS

A. Factor Extraction

For the recognition data, scree plots, VSS, and parallel analysis did not give conclusive results about the number of factors to extract, ranging from a wholly unrealistic 57 (scree plot) to a more realistic 4 (VSS). A model with four factors also resulted in the best model fit based on the AIC_c values, explaining 47 percent of the variance. For both correct and incorrect verifications, there were again inconclusive results, but in both cases a one-factor model lead to the best possible fit based on AIC_c. This explained 9 percent of variance for correct and 14 percent for incorrect continuations.

B. Recognition Model

Factor 1, the Gender factor, of the recognition model explained 18 percent of the variance. The positively loading song fragments, of which the most significant appear in Table 1, had a female vocal prominence in common and were released around the same time, between 2006 and 2013. There was a small correlation with being female (r = .14), daily listening hours (r = .14) and annual music events (r = .12), but not with any of the other demographic variables. Table 2 shows intensity and tonal conventionality as negative predictors and recurrence as a positive predictor, together explaining 11 percent of the variation.

Factor 2, the Age factor, explained 15 percent of the recognition variance. All the positive loading songs were released in the 1990s, whereas the low loading songs were released more recently (Table 1). Consistent with this pattern, there was a moderate correlation with age (r = .41). There was also a positive correlation with education (r = .32) and the frequency of attendance at music events (r = .22), also consistent with the profile of a (relatively) older participant. In addition, there were small correlations with a preference for world music (r = .10), "other" music (i.e., a genre that was not listed, r = .13), being a bowed string player (r = –.10) or a brass player (r = .11). A regression model with rhythmic irregularity, rhythmic conventionality, and event sparsity as positive predictors explained 11 percent of the variance, (Table 2).

Factor 3, the Rhythm factor, explained 8 percent of the variance in the recognition data. The positively loading song fragments all featured an element of rap or a prominent beat and all had male vocalists. Katy Perry's low-loading item did not have a prominent beat, but rather a female melodic vocal prominence (Table 1). The factor showed no correlation with

Table 1. Selection of the factor analysis results: Each test item with its starting point in the song and factor loading (i.e., *a*-value).

Recognition Factor 1: Gender

Artist – Song	Time (s)	Factor Loading
Fergie – Big Girls Don't Cry	231.0	.75
Katy Perry – Roar	49.4	.74
Leona Lewis – Bleeding Love	99.9	.74
Lou Bega – Mambo No. 5 (A Little Bit of …)	18.4	-.08

Recognition Factor 2: Age

Artist – Song	Time (s)	Factor Loading
The Fugees – Killing Me Softly	11.9	.80
Bryan Adams – (Everything I Do) I Do It for You	223.4	.75
Aerosmith – I Don't Wanna Miss a Thing	103.8	.73
Usher – Yeah!	85.7	.05

Recognition Factor 3: Rhythm

Artist – Song	Time (s)	Factor Loading
Eminem – Without Me	186.2	.90
Michael Jackson – Beat It	88.5	.67
Survivor – Eye of the Tiger	10.9	.61
Katy Perry – Roar	113.4	-.53

Recognition Factor 4: Residual

Artist – Song	Time (s)	Factor Loading
Las Ketchup – Ketchup Song	35.6	.89
Michael Jackson – Billy Jean	165.2	.80
Shakira – Hips Don't Lie	79.2	.36
The Human League – Don't You Want Me	77.6	-.43

Verification – Correct: Sing-along

Artist – Song	Time (s)	Factor Loading
Bill Medley & Jennifer Warnes – I've Had the Time of My Life	8.9	.93
Survivor – Eye of the Tiger	124.2	.79
UB40 – (I Can't Help) Falling in Love with You	37.4	.68
Survivor – Eye of the Tiger	168.1	-.50

Verification – Incorrect: Change

Artist – Song	Time (s)	Factor Loading
Aerosmith – I Don't Wanna Miss a Thing	32.5	.70
Sean King – Beautiful Girls	16.7	.64
The Fugees – Killing Me Softly	160.1	.64
Beyoncé – Crazy in Love	49.1	.06

Table 2. Estimated coefficients ($\hat{\beta}$) and standard errors of the highest-ranking regression models, predicting discrimination in the recognition and verification models.

	Rec. F1	**Rec. F2**	**Rec. F3**
Intercept	0.70 (0.08)	-0.21 (0.08)	-0.67 (0.06)
Intensity	-0.26 (0.07)		
Event Sparsity		0.19 (0.08)	
Recurrence	0.15 (0.07)		
Tonal Conventionality	-0.15 (0.06)		
Rhythmic Conventionality		0.20 (0.08)	-0.13 (0.06)
Rhythmic Irregularity		0.30 (0.09)	
Melodic Complexity			-0.21 (0.06)
Harmonic Complexity			-0.11 (0.05)

	Rec. F4	**Ver. Cor.**	**Ver. Inc.**
Intercept	-0.37 (0.11)	0.44 (0.04)	0.65 (0.02)
Intensity	-0.25 (0.22)		
Event Sparsity			-0.06 (0.02)
Recurrence	-0.21 (0.19)		
Tonal Conventionality			
Rhythmic Conventionality		-0.09 (0.05)	
Rhythmic Irregularity			
Melodic Complexity			
Harmonic Complexity			

Note. Standard error for each coefficient indicated in parentheses.

participants' gender, but did show correlations with age (r = .24), education (r = .17), and frequency of attendance at music events (r = .17), albeit smaller than in Factor 2. There were also small correlations with a preference for electronic music (r = .14), world music (r = .11), and "other" music (r = .18, and note that rap or hip-hop music was not given as an option). There were also small negative correlations with being a bowed string player (r = –.11) or woodwind player (r = –.13). The regression model on audio features explained 12 percent of the variance, with melodic complexity, rhythmic conventionality, and harmonic complexity as negative predictors (Table 2). Harmonic complexity did not have a significant contribution with a forced entry regression analysis (p = .08), but did appear consistently in the five highest ranked models with a coefficient magnitude bigger than the other coefficients, convincing us that it was also a potentially important predictor in our model.

Factor 4, the Residual factor, was the least explanatory factor of the model, explaining 6 percent of the variance. For this factor, it was also not clear what the positive loading songs had in common (Table 1). The factor was correlated with age (r = .11), education (r = .15), annual music events (r = .16), and being a brass player (r = -.16). The regression model explained only 4 percent of the variance in the factor loadings, with intensity and recurrence as negative predictors. The two predictors were not significant in a forced entry regression (intensity, p = .08; recurrence, p = .05), but both appeared consistently in the five highest-ranking regression models (Table 2).

C. Verification Models

The positive loading songs for the correct continuations, a selection of which are shown in Table 1, all appeared to be easy to sing along with. Thus, we coined this the Sing-Along factor. What stood out is that Survivor's 'Eye of the Tiger' has a fragment that loaded positively and a fragment that loaded negatively on the factor. In the song, the positively loading fragment preceded the negatively loading fragment. They both featured – almost – the same melody. The negatively loading fragment however, had timing slightly deviating from the first time the melody appeared, thus tricking the player that its continuation was incorrect whereas it actually was correct. The only demographic that the factor correlated with was being a singer (r = .12). There were also small correlations with the first three recognitions factors: Gender, Age, and Rhythm (F1: r = .20; F2: r = .14; F3: r = .17), suggesting that there was at least a small relationship between recognizing a song and verifying its continuation, for the first three factors. Finally, regression analysis showed that discriminability was predicted by rhythmic conventionality (Table 2), although this only explained 2% of the variance.

For the incorrect continuations, there were no song fragments with a negative loading. After a careful listen to the positively loading fragments, all the items seemed to have a very obvious change in the music 20–30 seconds into the fragment (Table 1). Thus, if a song continued incorrectly, 15 seconds after its correct continuation, the song would continue in a section of the song that sounded so different that its continuation would be obviously wrong. We called this the Change factor. It also showed a small correlation with being a singer (r = .13) and no correlation with other demographic variables. It showed small correlations with the first three recognition factors as well, Gender, Age, and Rhythm, in a similar pattern as the correct continuations (F1: r = .20; F2: r = .14; F3: r = .19). The correct and incorrect verification models, however, only showed a very small correlation with each other (r = .09). Finally, event sparsity was a significant predictor for discriminability (Table 2), explaining only 3% of the variability.

D. Summary of Findings

Our factor analysis resulted in a four-factor recognition model and a single-factor verification model for both correct and incorrect continuations. The Gender factor seemed to represent a preference, and thus better recognition, for music with a high female vocal prominence. This was corroborated by a small correlation with gender. The Age recognition factor appeared to be a factor of age, where people only recognized music from a certain time period. This can be related to the reminiscence bump (Krumhansl & Zupnick, 2013; Schulkind et al., 1999). The Rhythm factor appeared to represent a preference for music that falls under the genre of rap/hip-hop, but could also be interpreted as a preference for or specific attention to music's intense and strongly rhythmic musical characteristics. The Residual recognition factor was the only factor that wasn't easily retraceable. We need more detailed and different demographic information and a more careful analysis of the music fragments to uncover this last factor.

The verification with correct continuation model (Sing-along factor) appeared to be a representation of exactly what we asked the players to do; sing along with the music. The easier a fragment was to sing along to, and the better participants were at singing along, the better they were at recognizing that the song continued in the correct spot. The incorrect continuation model (Change factor) seemed to also be a result of the task at hand, instead of individual differences in long-term musical memory. It likely represented a technique that people used; listening whether the overall sound of the song was very different pre- and post-mute. Those who did not or could not use this technique, had a lot of difficulty recognizing that the song continued in the incorrect spot.

IV. DISCUSSION

The aim of this study was twofold. First, we aimed to uncover individual differences in long-term musical memory based on the large dataset from Hooked on Music. Second, we wanted to demonstrate a technique that had not previously been used in musical memory research before; MIRT.

Based on existing literature, we hypothesized about several sources of individual differences; age, music preference, level of processing, and attention to specific musical characteristics. The results of our exploratory factor analysis did show some effects of age and music preference or attention to specific musical characteristics. In the recognition model, we found factors that could be related to preference for female vocals (Gender factor), music from a specific genre or with specific musical characteristics (Rhythm factor), and age (Age factor).

The verification models did not show many factors of individual differences, except what we interpret as a player's ability to execute the task at hand. For verification items that continued correctly, the ability to sing along with the music (correlated with being a singer) determined individual differences. For verification items that continued incorrectly, arguably a much harder task, it appeared to be the strategy to listen for obvious changes post-mute compared to pre-mute, that caused individual differences. This verification task was implemented to verify that players actually recognized the song, which the task does well given the correlation with all the recognition factors. However, it did not seem to add much to the question of what causes individual differences in musical memory.

A few future directions are of interest for the Hooked on Music project and memory research in general. For example, our current participants were asked quite few and basic demographic questions. The motivation for this was that the more instructions and reading material we presented our players with, the fewer people were encouraged to participate. In the current format, only 12% filled in the complete demographic questionnaire. Nevertheless, more detailed demographic information could help us learn more about the results of the factor analysis; how do they relate to, e.g., to the *MUSIC* model on musical preferences (Rentfrow et al., 2011), or the Gold-MSI (Müllensiefen et al., 2014)? A possible way to motivate the players to fill out these questions would be to provide them with some form of feedback, e.g., their Gold-MSI score, without influencing their performance.

Another focus for future research would be the individual differences in attention that people pay to specific musical characteristics. One of our recognition factors could be the result of this, but could also be the result of a music genre preference. The research on attention that people pay to different musical characteristics, or their sensitivity to it, is lacking, with only a few studies on attention to the beat (Grahn & McAuley, 2009; McAuley et al., 2006). More research should be done on individual differences in the attention for specific music characteristics, before we can theorize about its effects on our memory.

A final future direction would be the area of dementia research and intervention. Firstly, a version of Hooked on Music might be adapted to be used by dementia patients, enabling them to reminisce with music and benefit from its possible positive effects (McDermott, Orrell, & Ridder, 2014). Other applications have already shown that dementia patients are able to use modern technology (Alm et al., 2005). Secondly, in addition to the possible practical use, a dementia adaption of Hooked on Music could be linked to individual differences research. On the one hand, our current knowledge could provide better personalized playlists for dementia patients. On the other hand, data from dementia patients could reveal more about individual differences in memory research and what is impaired with people suffering from dementia.

To conclude, this study has taken a next step in uncovering musical memory, by demonstrating the use of MIRT analysis and taking an individual differences approach. It has shown that age and musical preferences or musical characteristics are important determinants of individual differences in long-term musical memory. This can be of use for researchers studying musical memory and developing computations offering personalized playlists, to both the general public and a clinical population. Future directions with a similar citizen science project can also study memory for other music genres besides pop-music, or study cross-cultural differences in long-term musical memory. In addition to providing new insights on long-term musical memory, we hope this study encourages future studies to apply similar research techniques.

ACKNOWLEDGMENT

HH is supported by a Horizon grant (317-70-10) of the Netherlands Organization for Scientific Research (NWO). JAB is supported by a Talent Grant from the Amsterdam Brain and Cognition centre. Hooked on Music was created in collaboration with the Museum of Science and Industry in Manchester, UK.

REFERENCES

Alm, N., Dye, R., Astell, A., Ellis, M., Gowans, G., & Campbell, J. (2005). Making software accessible to people with severe memory deficits. In *Proceedings of the 2005 International Conference on Accessible Design in the Digital World*, 16.

Burgoyne, J. A., Balen, J. Van, Bountouridis, D., & Honing, H. (2013). Hooked: A game for discovering what makes music catchy. In *Proceedings of the International Society for Music Information Retrieval (ISMIR)*, 245-250.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671-684.

Grahn, J. A., & McAuley, J. D. (2009). Neural bases of individual differences in beat perception. *NeuroImage*, *47*(4), 1894-1903.

Hyde, T. S., & Jenkins, J. J. (1973). Recall for words as a function of semantic, graphic, and syntactic orienting tasks. *Journal of Verbal Learning and Verbal Behavior*, *12*(5), 471-480.

Krumhansl, C. L., & Zupnick, J. A. (2013). Cascading reminiscence bumps in popular music. *Psychological Science*, *24*(10), 2057–2068.

McAuley, J. D., Frater, D., Janke, K., & Miller, N. S. (2006). Detecting changes in timing: Evidences for two modes of listening. IN *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC)*, 188-189.

McDermott, O., Orrell, M., & Ridder, H. M. (2014). The importance of music for people with dementia: The perspectives of people with dementia, family carers, staff and music therapists. *Aging & Mental Health*, *18*(6), 706-716.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688-691.

Rentfrow, P. J., Goldberg, L. R., & Levitin, D. J. (2011). The structure of musical preferences: A five-factor model. *Journal of Personality and Social Psychology*, *100*(6), 1139-1157.

Schulkind, M. D., Hennis, L. K., & Rubin, D. C. (1999). Music, emotion, and autobiographical memory: They're playing your song. *Memory & Cognition*, *27*(6), 948-955.

Ter Bogt, T. F. M., Mulder, J., Raaijmakers, Q. A. W., & Nic Gabhainn, S. (2010). Moved by music: A typology of music listeners. *Psychology of Music*, *39*(2), 147-163.

Van Balen, J., Burgoyne, J. A., Bountouridis, D., Müllensiefen, D., & Veltkamp, R. C. (2015). Corpus analysis tools for computational hook discovery. In *16th International Society for Music Information Retrieval Conference (ISMIR)*, 227-233.

Van Balen, J., Burgoyne, J. A., Wiering, F., & Veltkamp, R. C. (2013). An analysis of chorus features in popular song. In *Proceedings of the 14th International Society for Music Information Retrieval Conference (ISMIR)*.