Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009) Jyväskylä, Finland
Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, Päivi-Sisko Eerola (Editors)

Music Preferences Based on Audio Features and its Relation to Personality

Greg Dunn

User Experiences Group, Philips Research Europe, The Netherlands
peter.gregory.dunn@gmail.com

ABSTRACT

Recent studies have summarized reported music preferences by genre into four broadly defined categories, which relate to various personality characteristics. Other research has indicated that genre classification is ambiguous and inconsistent. This ambiguity suggests that research relating personality to music preferences based on genre could benefit from a more objective definition of music. This problem is addressed by investigating how music preferences linked to objective audio features relate to personality. Participants (N = 354; 165 males) completed a personality measure and rated their preference to 120 music clips from various music (sub-)genres (e.g., classical, rock). Principle Components Analysis revealed a nine-component model that accounted for 61% of the variance in music preference ratings. Audio features computationally derived from the music clips were subsequently analysed to discriminate between music contained within each of these factors. In addition, participants' estimated music preference scores to these factors were related to personality facets. Aggregated results showed, for example, that Excitement-Seeking was positively related to music with a greater number of percussive events while negatively related to music with fewer percussive events. Results are discussed in terms of how objective features can provide greater insight into how music preferences relate to personality.

I. INTRODUCTION

It has been suggested that music acts as a mirror to reflect our own inner self (DeNora, as cited in Juslin & Sloboda, 2008). Perhaps this is best demonstrated by recent research that has related music preferences to personality (e.g., Delsing, Ter Bogt, Engels, & Meeus, 2008; George, Stickle, Rachid, & Wopnford, 2007; Rentfrow & Gosling, 2003; Zweigenhaft, 2008). These studies related participants' music preferences, reported by genre, to various personality dimensions. Pachet and Cazaly (as cited in Aucouturier & Pachet, 2003), however, have shown that genre labels are ambiguous and inconsistently used. Experience also tells us that most individuals actively listen to a range of music that would be labelled under a variety of different genres, and reported preference to any given genre does not mean that one prefers all music labelled under that genre. So, relations reported between music preferences and personality remains somewhat vague with respect to what preferences to a specific genre really mean. This issue has been explored in the present paper by showing how specific audio features extracted from music can distinguish between music preferences, which have been related to personality. In doing so, it is argued that the present paper provides greater insight into how specific audio features found in music might reflect aspects of our personality.

Having stated the main objective of the present paper, the remainder of the Introduction section outlines previous research relating music preferences to personality, and introduces how audio feature extraction might benefit this research. The aims and hypotheses of the present study then conclude this section.

A. Personality and Music Preferences

Early research examining the relation between personality and music preferences has varied somewhat with respect to how both personality and music preferences were measured (e.g., Arnett, 1992; Cattell & Anderson, 1953; Cattell & Saunders, 1954; Litle & Zuckerman, 1986; McCown, Keiser, Mulhearn, & Williamson, 1997; McNamara & Ballard, 1999; Schwartz & Fouts, 2003). For instance, Cattell and his colleagues (1953; 1954) measured participants' music preferences using specific music recordings while other researchers used reported genre preferences to measure music preferences (e.g., Arnett; Litle & Zuckerman; McNamara & Ballard). In addition, five different personality measures had been used amongst this research. As a result, these differences in measured personality and music preferences made direct comparisons regarding the outcomes between these studies somewhat difficult.

Since 2003, however, research relating personality and music preferences seemed to have aligned with respect to how both personality and music preferences were measured. This aligning was arguably due to a paper by Rentfrow and Gosling (2003). In this paper, Rentfrow and Gosling described their research where they developed, confirmed, and validated their own measure of reported music genre preferences and related these preferences to the Big Five personality dimensions. Like Rentfrow and Gosling, later research relating music preferences to personality measured reported music preferences according to genre and related these preferences to the Big Five (e.g., Delsing et al., 2008; George et al., 2007; Zweigenhaft, 2008). The genre categories used among this literature has been highly similar and included genres such as Blues, Classical, Country, Dance, Heavy Metal, Jazz, Pop, Rap, Rock, R&B. Participants in each of these studies rated their preference to these genres and other genres on either 5-point or 7-point Likert scales. Afterward, these preferences were often grouped according to preference ratings and then related to the Big Five.

Within trait theories of personality, the Big Five has been arguably the most accepted model of personality that currently exists (John & Srivastava, 1999). As its name suggests, the Big Five measures five trait dimensions that have been identified and described by Costa and McCrae (1992) as:

1. *Neuroticism* (N) – individual's propensity to feel fear, sadness, anger, and other emotions of negative affect.
2. *Extraversion* (E) – individual's propensity to be sociable, assertive, active, and prefer exciting environments.
3. *Openness to Experience* (O) – individual's propensity toward intellectual curiosity, imagination, and originality.
4. *Agreeableness* (A) – individual's propensity toward being altruistic, helpful, and empathetic toward others.
5. *Conscientiousness* (C) – individual's propensity toward cleanliness, orderliness, determination, and self-control.

From this review, it appears that the Big Five is prominent among the trait models of personality and among the recent research that has related personality to music preferences. Given this prominence, it would seem logical to continue using some form of measurement of the Big Five for future studies to be comparable with previous research that has investigated the relation between personality and music preferences. To further support comparisons to this previous research, it is important to consider genre labels often used to measure music preference. Nonetheless, the next section will argue why genre labels might be too vague to give an accurate measure of music preferences.

B. Genres and Audio Feature Extraction

Much of the research that has investigated music preferences and its relation to personality has asked participants to rate their music preference using given genre labels (e.g., Arnett, 1992; Delsing et al., 2008; George et al., 2007; Litle & Zuckerman, 1986; McNamara & Ballard, 1999; Rentfrow & Gosling, 2003; Zweigenhaft, 2008). At first, this appears to make sense; individuals often arbitrarily use genre labels in conversation to describe their music preferences (Rentfrow & Gosling, 2006). These labels are also often used in various areas of research to describe music or music preferences (e.g., Juslin & Sloboda, 2008; Levitin, 2006; North & Hargreaves, 1999). Simply put, genre labels are a convenient and effective way to describe various styles of music playing on the whole.

Despite the convenience and common use of genre labels, there are several reasons to suggest that genre labels might not be the most accurate method to measure music preferences. First, sociologists have pointed out that genre labels are used as a tool by the music industry as part of its strategy to sell music to various audiences (e.g., Negus, 1996). This would suggest that genre labels are ultimately somewhat subjective in nature. Perhaps it is this reason that people have often used several genre labels to describe their musical taste, which leads to the second reason why genre labels might have limited accuracy when measuring music preferences.

Second, our everyday experience suggests that genre labels are neither fully descriptive of individuals' music preferences, nor are these labels intended to mean that all music contained therein would be similarly enjoyed by a given individual. For example, an individual, let us call her Katrina, may describe her music preferences to include music from Rock, Heavy Metal, and Blues genres. And while Katrina might love the Rock band, U2, she might abhor another Rock band like Coldplay. Granted, there could be many inter-related reasons for Katrina to like one Rock band (in this case, U2), and not like another (Coldplay). These reasons could include, but are not limited to:

- how a given music artist or group has been marketed by the music industry (Negus, 1996);
- the social status attached to given music artists and groups (North & Hargreaves, 1999; Rentfrow & Gosling, 2006);
- the emotional response that can be instigated by certain music or songs (Juslin & Sloboda, 2008);
- (emotionally significant) memories that have been linked to certain songs (Levitin, 2006).

Despite these reasons, however, previous research has shown that genre labels have been sufficiently accurate to show measureable trends in music preferences (e.g., Delsing et al., 2008; George et al., 2007; Rentfrow & Gosling, 2003). Using our example again, Katrina has provided a measureable trend that indicates she likes Rock, Heavy Metal, and Blues music. This description is accurate to a point, but perhaps greater descriptive accuracy would be gained if there were specific audio features that were common among all the music that Katrina listens to within those three genres. This argument leads to the third and final reason why genre labels might have limited accuracy when measuring music preferences.

Third, certain audio features might give greater descriptive accuracy of individuals' music preferences compared to genre labels. As stated earlier in this Introduction, Pachet and Cazaly (as cited in Aucouturier & Pachet, 2003) conducted a study that showed different genre taxonomies employed by various music resources representing the music industry were inconsistent in their approach to music categorization. This finding prompted Aucouturier and Pachet to investigate other approaches to music categorization, which included audio feature extraction.

Audio features contained in music have been shown to communicate emotion to the listener (e.g., Juslin, 2000), and the relation between music and emotion has been intensely studied (see Juslin & Sloboda, 2008). Consider that personality traits, such as the Big Five, predispose individuals to experiencing certain emotional states (Rusting, 1998). Given this, it should not be surprising that individuals with specified personality traits are shown to have preferences to music with an empirically-defined set of audio features. Nevertheless, there has been no known research published, which has investigated how personality relates to music preferences, and how audio features might distinguish these music preferences. This paper introduces a first foray into this area of research, and shows an approach to investigating the relation between personality and music preferences using audio feature extraction.

C. Aims and Hypotheses

Briefly, three primary aims were specified for this study:

1. Develop a preliminary model of music preferences using audio stimuli (i.e., specific songs).
2. Relate these music preferences to personality traits.
3. Discriminate between music preferences using extracted audio features.

To make results comparable to research that has investigated the relation between personality and reported music preferences according to genre, songs were sampled from several genres previously used to represent music preferences (e.g., Rentfrow & Gosling, 2003). How these songs were sampled has been further described in the Method section.

Given that this study was an initial step toward linking personality, music preferences, and extracted audio features, there were few testable hypotheses that were specified. The first hypothesis was concerned with the impact that familiarity has on preferences toward specific songs. Common sense would suggest that individuals are more likely to prefer songs that they also report as being familiar to them. The second hypothesis has taken into account the literature that has indicated the effects of age and gender on music preferences (e.g., Arnett, 1992; George et al., 2007; McCown et al., 1997;

McNamara & Ballard, 1999; Zweigenhaft, 2008). Similarly, the third hypothesis has taken into account how music training affects music perception (e.g., Lamont & Dibben, 2001; Levitin, 2006; McAdams & Matzkin, 2001), which might also affect music preference. These hypotheses were not central to the aims of the present study, but instead were either of secondary interest (familiarity) or may have impacted how a model of music preferences using audio extracted features would be developed (age, gender, music training). Therefore, literature related to these hypotheses was not specifically addressed in the Introduction. Before moving on to the Method, the hypotheses for the present study may be summarized as:

- H1. Familiarity toward a given song will be positively related participants' song preference ratings.
- H2. Age and gender will show significant interaction effects on participants' music preference.
- H3. Music training will show a significant interaction effect on participants' music preference.

Reference to these aims and hypotheses will be done when analyzing the results and discussing the outcomes of this study.

II. METHOD

The experiment described in this paper was part of a larger experiment investigating the relations between personality and music preferences. The following sections describe only the parts of this experiment that are considered to be relevant to the aims and hypotheses previously stated in the Introduction.

A. Participants

Participants (N = 354; 165 males) volunteered in response to recruitment announcements provided over the Internet via several means (e.g., mailing lists, forums). Most participants reported having American nationality (n = 153), followed by Canadian (n = 64), British (n = 31), and various other nationalities from around the world (n = 106). Participants' ages ranged from 18 to 68 years (M = 31.52, SD =11.02).

B. Materials

Participants listened to 120 different music clips streamed over the Internet and played from their own computer. Each music clip lasted 20 seconds taken from what was believed to be the most representative portion of the entire music recording (i.e., song). Based on the genres used by Rentfrow and Gosling (2003), these music clips ranged across ten different genres: Blues, Classical, Country, Dance, Heavy Metal, Jazz, Pop, Rap, Rock, R&B (12 clips per genre). Music clips were classified by genre based on converging information from three different music industry sources (Amazon.com, 2007; AMG, 2007; Last.fm, 2007). No music recording was represented twice in different music clips.

Participants used an Internet interface provided with each music clip to answer the following items with each song (the Likert-scale anchors are provided in brackets):

1. In your opinion, how much do you like this song? (1 = Strongly Dislike, 2 = Dislike, 3 = Neutral, 4 = Like, 5 = Strongly Like)
2. Are you familiar with this song? (1 = Not at all, 2 = Maybe a little, 3 = I know I've heard it before, 4 = I'm very familiar with the song, 5 = I'm a big fan)
3. In your opinion, would you like to have this song and songs similar to this (from the same artist, etc.) recommended to you in the future? (1 = Certainly not, 2 = Unlikely, 3 = Maybe, I don't care either way, 4 = Probably, 5 = Definitely)
4. Would you consider adding this song to your music collection (e.g., any form of downloading, CD purchase)? (1 = Never, 2 = Unlikely, 3 = Maybe, I don't care either way, 4 = Probably, 5 = Definitely or already have it in my collection)

Questions 1, 3, and 4 were used as a measure of participants' music preference per song (Cronbach's alpha = .95). Along with giving demographic information (age, gender, nationality, years of music training, and hours per week listening to music), participants were asked to fill out the following questionnaire:

- *Revised NEO Personality Inventory* (NEO PI-R): is used to measure participants' personality (Costa & McCrae, 1992). Participants rated 240 items on a scale from 1 (Strongly Disagree) to 5 (Strongly Agree), which loaded onto the Big Five personality dimensions. This measure gives aggregated scores for the five dimensions, as well as the six facets contained within each dimension.

C. Procedure

Participants were tested over the Internet. After viewing an Informed Consent page, participants began the experiment by providing their demographic information. Then approximately half of the participants completed the NEO PI-R first and half listened and responded to the questions about the music clips first. Participants who completed the NEO PI-R first then listened and responded to the questions about the music clips. Those who completed the music portion first then completed the NEO PI-R. The music clips were given in a counterbalanced Latin-square design. Lastly, participants were provided with a debriefing screen and their own NEO PI-R personality report.

III. RESULTS

A. Familiarity to Music

The first hypothesis addressed if familiarity was related to participants' preference ratings to the same song. To express this relation, one correlation coefficient was calculated between participants' familiarity and their preference ratings toward the same music clip. The ensuing calculation was $r = .62$, $p < .001$, indicating that participants' familiarity was related to their preference score for the same song. Still, this relation was not considered to negatively impact further analyses because the relation between song familiarity and preference is assumed to be an integral part of music preference. That is, we tend to listen more often to songs we like.

B. Music Preference Components

Having addressed the first hypothesis, the next task switched to addressing the first aim specified for this paper, which was to develop a preliminary model of music preferences using audio stimuli. To address this aim, a Principal Components Analysis (PCA) was done to find the exploratory dimensions of music preferences among the 120 music clips rated by participants. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO statistic) for this analysis was .91, indicating strong patterns of music preference scores toward the music clips. Due to the large number of data points, however, there was an imminent danger of over-extracting the number of components. Overextraction could have negative implications when trying to confirm the model in future experimental samples (Zwick & Velicer, 1986). To avoid overextraction, several criteria were used to decide how many components would be retained (see Floyd & Widaman, 1995): the Kaiser-Guttman criterion (i.e., eigenvalues equal to or greater than 1), a scree test, and the interpretability of the component loadings for a given solution.

Using these criteria, a 9-component PCA solution was found, which accounted for approximately 61% of the variance in participants' music preference scores toward the music clips. For instance, Figure 1 shows the scree plot, which indicates a marked division between the 9th and 10th components in the bend or elbow of the scree curve. This criterion also meant that the minimum eigenvalues in this solution were all above 2.

Rotations on the 9-component solution were done to increase the interpretability of this solution and facilitate participants' component score estimates (Nunnally, 1967). Varimax rotation is the most commonly used method (Floyd & Widaman), which is further evidenced by previous work in psychology that has developed measures of personality (e.g., Costa & McCrae, 1992), and of music preferences (e.g., Rentfrow & Gosling, 2003). Varimax rotation is an orthogonal method that attempts to maximize the variance of the loading values within each component (Nunnally). Nonetheless, despite the popular use of Varimax rotation, it was assumed that participants' music preferences toward one musical style could potentially be correlated with another style. This justified implementing an oblique rotation, which in this case was Promax rotation.

Briefly, Promax rotation is a two-step process, which begins by obtaining a Varimax-rotated solution (Tabachnick & Fidell, 2007). This solution is then modified to an oblique rotation that reduces low loadings to near-zero values. In this way, the contrast between high and low loadings is improved. Regardless of the rotation, both the Varimax and Promax rotation solutions provided highly similar results. Nonetheless, the Promax rotation left open the possibility for correlations among components, and did provide a rotation solution that improved the distinction between high and low component loadings within each component. For these reasons, Promax rotation was used to communicate the pattern matrix for the final 9-component solution. Table 1 presents an abbreviated pattern matrix, which provides component loadings of |.400| or higher for a limited number of clips. The first two columns in Table 1 list the music represented by clips given to participants in this experiment, while the subsequent nine columns provide the component loadings for these clips in each of the nine components. These component loadings express the correlation between preference scores toward a given song and extracted scores for a given component (Tabachnick & Fidell). Complete tables for both the structure and pattern matrix with all loading values may be provided by request to the author.

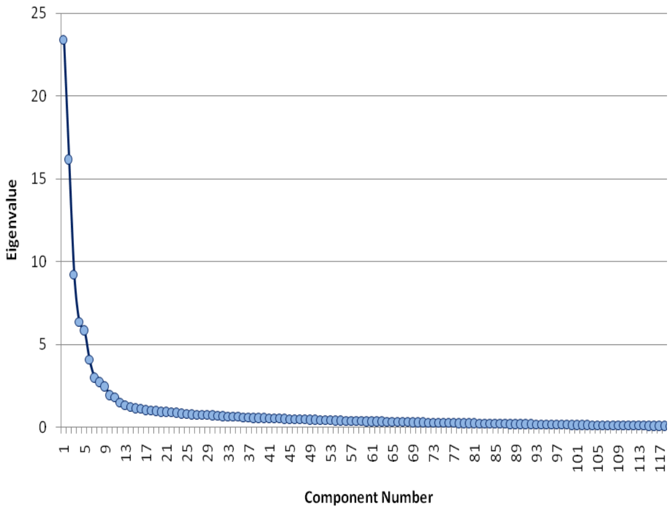

Figure 1. Scree plot indicating the Eigenvalues (y-axis) for each of the potential 120 Component Numbers (x-axis). The scree curve indicates a marked division between the 9th and 10th component.

C. Relating Music Preferences to Personality

Once the Principal Components Analysis (PCA) had been used to identify the music preferences components (MPCs), the second specified aim was to relate participants' predicted scores for each of these MPCs to their personality facet scores. As mentioned in the Method section, personality facets are more detailed aspects of each of the Big Five dimensions: Neuroticism (N), Extraversion (E), Openness (O), Agreeableness (A), and Conscientiousness (C; Costa & McCrae, 1992).

Before investigating the relation between participants' predicted music preferences and their personality, it was necessary to ensure that this relation would not be biased by other factors, such as gender, age, and music training. So, to see how these other factors influenced participants' predicted scores to the nine MPCs, a 2 (gender) x 9 (component) mixed ANCOVA was done with participants' predicted component scores as the DV, and age and years of music training as covariates. By doing so, this investigation addressed the second and third hypotheses listed in the Introduction. Within-subjects effects showed that predicted scores toward each of the nine MPCs had interaction effects with gender ($F = 10.53$ (8, 2,800), $p < .001$, partial $\eta^2 = .03$), age ($F = 21.17$ (8, 2,800), $p < .001$, partial $\eta^2 = .06$), and music training ($F = 12.66$ (8, 2,800), $p < .001$, partial $\eta^2 = .04$). Given these results, variance in participants' predicted music preference scores accounted for by gender, age, and music training was partitioned out. Thus, further statistical testing used participants' residual component scores after these three variables had been partitioned out.

Using these residual scores, nine stepwise regressions were done to ascertain predictive equations for participants' preferences toward the nine MPCs given their personality facet scores. Table 2 gives the standardized regression coefficients (β) per MPC given personality facets. In Table 2, the top row of numbers identifies each of the MPCs, along with its associated multiple regression coefficient of determination (R^2). The personality facets found to significantly predict a proportion of variance in any one of these MPCs are listed down the first column. Cells provide the standardized regression coefficients

Table 1. Pattern Matrix Factor Loadings for the 9-Component Promax-Rotated PCA Solution.

Music Clip Title	Artist/Composer	Music Preference Component Number								
		1	2	3	4	5	6	7	8	9
Dangerous	Busta Rhymes	***.907***								
Dirt Off Your Shoulder	Jay Z	***.895***								
Country Grammar (Hot Shit)	Nelly	***.884***								
Lester Swings	Lester Young		***.797***							
Locomotion	John Coltrane		***.788***							
All Blues	Miles Davis		***.777***							
Fall at Your Feet	Crowded House			***.784***						
Back For Good	Take That			***.750***						
Life for Rent	Dido			***.733***						
Canto Della Terra	Andrea Bocelli			***.549***		*.487*				
Baby One More Time	Britney Spears	*.430*		***.509***						
Back in Black	AC/DC				***.783***					
Paranoid	Black Sabbath				***.756***					
Smoke on the Water	Deep Purple				***.755***					
Nookie	Limp Bizkit	*.442*			***.518***					
Till Eulenspiegels lustige Streiche	Richard Strauss					***.856***				
Symphony No. 3, "Eroica"	Ludwig van Beethoven					***.845***				
Piano Concerto No. 1	Peter Ilitch Tchaikovsky					***.840***				
My Heart Skips a Beat	Buck Owens						***.809***			
It Wasn't God Who Made Honky Tonk Angels	Kitty Wells						***.807***			
Don't Rock the Jukebox	Alan Jackson						***.786***			
All the Kings Horses	Luther Allison							***.778***		
I Smell Trouble	Johnny Winter							***.777***		
Nice Problem to Have	The Jeff Healey Band							***.727***		
Pitiful	Big Maybelle		*.473*					***.552***		
You Keep Me Hangin' On	Diana Ross & the Supremes								***.659***	
Space Oddity	David Bowie								***.623***	
Something	The Beatles								***.606***	
Jumpin Jack Flash	The Rolling Stones				*.462*				***.543***	
I Walk the Line	Johnny Cash						*.478*		***.499***	
Talisman	Air									***.730***
It Began in Africa	The Chemical Brothers									***.701***
Push Upstairs	Underworld									***.662***
Destiny	Vanessa-Mae			*.506*						***.590***
Du Hast	Rammstein				*.434*					***.522***

Note. N = 354. All factor loadings |.400| or larger are provided in *italics*; the highest factor loadings within each component are in **bold**.

Table 2. Standardized Regression Coefficients (β) per Music Preference Component given Personality Facets.

Personality Facet	Music Preference Component Number 1	2	3	4	5	6	7	8	9
R^2=	.19	.11	.08	.09	.13	.07	.06	.10	.17
N1: Anxiety								.10	
N5: Impulsiveness							-.11		
E1: Warmth	.27		.17						
E4: Activity						-.14			
E5: Excitement-Seeking	.34			.26	-.12		.17	.13	
O1: Fantasy								.09	
O2: Aesthetics		.16	-.11	-.19	.22		.14		.17
O3: Feelings					-.11				
O4: Actions		.14	-.13						.16
O5: Ideas		.13		.15	.21				
O6: Values					-.18	-.20		.12	.15
A2: Straight-forwardness						-.15			
A3: Altruism	-.20						-.16		-.15
A6: Tender-Mindedness						.13			
C1: Competence								.22	
C2: Order	.13								
C4: Achievement Striving									.19
C5: Self-Discipline			.13					-.20	-.23

Note. N = 354. Cells represent standardized regression coefficients (β). All values shown are significant at $p < .05$.

(β) for the designated MPC given the designated personality facet. All R^2 and β values were significant at $p < .05$. To confirm these results, a randomly selected participant sub-sample was taken and nine regression analyses were done to provide the R^2 between participants' preference scores from this sub-sample, and their predicted scores for each of these predictive equations (Tabachnick & Fidell, 2007). These values were then compared to the R^2 values provided by the original predictive equations that are provided in Table 2. With exception to component 6, the R^2 values drawn from the sub-sample were significant and comparable to the R^2 values from the original analyses. This suggested that eight of the nine predictive equations were reliable estimates of music preferences given personality, with exception to the predictive equation for component 6.

D. Audio Feature Extraction

The final aim of this study was to discriminate between music preferences components (MPCs) by using extracted audio features from the music clips. To achieve this aim, a Multiple Discriminant Analysis (MDA) was done using the extracted audio features from the music clips as membership predictors for each of the nine MPCs identified by the Principle Components Analysis (PCA).

Before conducting the MDA, it was necessary to filter the music clips to be used in this analysis. As a minimum criterion, only clips that were clearly categorized into one of the PCA's nine MPCs were included in the MDA. Stated objectively, this criterion meant that music clips had to have a factor loading with a greater magnitude of |.400| on only one of the PCA components to be included in the MDA. As a result, 16 of the original 120 music clips did not attain this minimum criterion and were not included in this analysis.

Of the remaining 104 music clips, extracted audio features were taken from these clips approximately every 950 ms, resulting in 21 observations per clip. So, there were 2,184 data points used for each audio feature in the MDA. A total of 85 audio features were extracted from each music clip. These audio features are divided into four general categories: 1) spectral-temporal signal properties, 2) percussive event properties, 3) tonal properties, 4) rhythmic properties. Unfortunately, a precise description of the audio extraction software used in this analysis cannot be given because of the confidential nature of this software tool developed within Philips Research. Generally speaking, however, the nature of this tool is similar to other software dedicated to audio feature extraction (e.g., Lartillot & Toiviainen, 2007; Pampalk, 2004). Finally, all observations obtained from each of the 85 audio features were transformed into *z-scores* to ensure that these audio features were all comparable along the same scale.

A stepwise MDA eliminated 53 redundant audio features from the original 85 audio features extracted. This left 32 audio features that were used in eight functions to discriminate between the music contained in each of the nine MPCs. To evaluate the overall significance of the MDA, a chi-square test of Wilks' Lambda (Λ). The result was Λ = .009, χ^2 (256, N = 2, 184) = 10,081.43, $p < .001$, partial η^2 = .44 with 95% confidence limits from .40 to .46. Further tests of Λ indicated that each function also significantly added to the discriminant ability of the MDA. This was shown by a chi-square test of Λ for the last and weakest discriminant function, Λ = .845, χ^2 (25, N = 2, 184) = 364.20, $p < .001$, partial η^2 = .02 with 95% confidence limits from .00 to .02. Nonetheless, the first four discriminant functions accounted for a total of 85.5% of the between-group variability amongst the Music Preference Components (MPCs), with each function separately accounting for at least 10% of this total variance. For this reason, as well as consideration of space, only results for these four functions will be described in further detail.

A canonical R^2 was used for each of the four discriminant functions to express the relation between the extracted audio features and the MPCs. In order from the first to the fourth discriminant function, these canonical R^2 values were R^2 = .77,

R^2 = .61, R^2 = .51, R^2 = .45. Similar to the Λ values expressed above, each of these values were significant at p < .001. Figure 2 gives a visual representation of the first two discriminant functions that used extracted audio features to discriminate music clips according MPCs. Only these two functions were used because they provided the best visual distinction between the music contained in each of the nine MPCs. Furthermore, Table 3 provides the group means (i.e., group centroids) for each of the discriminant functions.

Using Figure 2 and Table 3, two extremes are seen from the means in the first discriminant function. The negative extreme contained music from component 1 (M = -2.42, SD = 0.76) and component 9 (M = -2.55, SD = 1.18). The positive extreme had music from component 2 (M = 1.56, SD = 1.19), component 6 (M = 1.07, SD = 0.85), and at its most extreme, component 5 (M = 3.76, SD = 1.40). Further tests indicated that the means between these extremes were significant (p < .001). Using the function loadings provided by the MDA, the best predictors for discriminating between these extreme groups included audio features related to percussive event properties. These loadings indicated that music from the negative extreme tended to have more percussive events with shorter periods of time between these events compared to music from the positive extreme.

Interpretation of the function loadings provided by the MDA indicated that percussive events were also important when discriminating between extremes along the second discriminant function. At its most negative extreme along this discriminant function was music from component 5 (M = -2.34, SD = 1.31), and from component 1 (M = -1.56, SD = 0.95). At its positive extreme was music from component 4 (M = 0.98, SD = 0.70), component 6 (M = 1.24, SD = 0.78), component 7 (M = 1.30, SD = 0.89), and component 8 (M = 1.05, SD = 0.96). Again, further tests indicated that the means between these two extremes were significant (p < .001). These loadings indicated that music taken from the negative extreme tended to have greater variation in the timing between percussive events compared to music from the positive extreme.

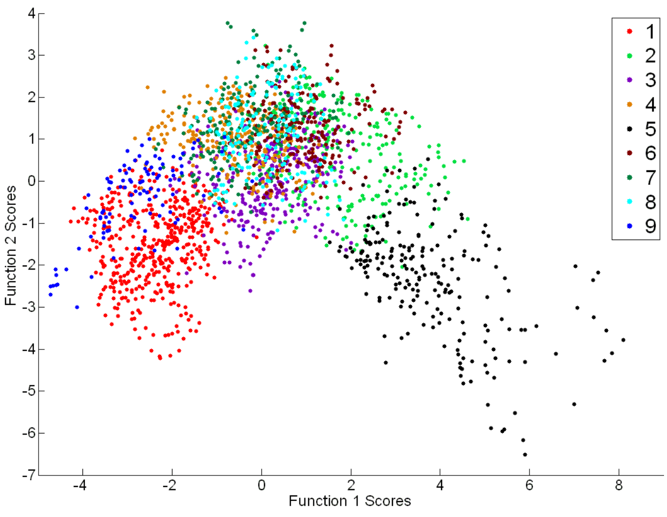

Figure 2. Visual representation of the first two discriminant functions that used extracted audio features to discriminate music clips according to MPCs.

Table 3. Means per Group (Centroids) taken from MDA.

Music Preference Component	Discriminant Function 1	2	3	4
1	-2.424	-1.562	-0.569	0.487
2	1.558	0.645	-1.515	0.491
3	0.157	0.017	1.517	1.102
4	-0.647	0.975	0.767	-1.325
5	3.756	-2.339	0.152	-0.929
6	1.065	1.235	0.781	0.763
7	-0.372	1.300	-1.233	-0.388
8	0.017	1.052	-0.320	0.021
9	-2.549	-0.274	0.607	-1.811

Note. N = 2,184. Cells represent group means (M).

For the third discriminant function, the negative extreme contained of music from component 2 (M = -1.51, SD = 1.07) and component 7 (M = -1.23, SD = 0.93). At the positive extreme was music from component 3 (M = 1.52, SD = 0.90). Further tests indicated that the means between these two extremes were significant (p < .001). The loadings provided by the MDA indicated that the best predictors for discriminating between these extreme groups included audio features related to tonal properties. These loadings indicated that music from the negative extreme tended to be played more often in minor key and also tended to have a more complex tonal structure compared to music from the positive extreme.

The negative extreme of the fourth discriminant function contained music from component 4 (M = -1.33, SD = 0.86) and component 9 (M = -1.81, SD = 0.82). At its positive extreme was music from component 3 (M = 1.10, SD = 0.84) and component 6 (M = 0.76, SD = 0.83). Further tests indicated that the means between these two extremes were significant (p < .001). Similar to the third discriminant function, the loadings provided for the fourth function showed that the best predictors for distinguishing between these extreme groups included audio features related to tonal complexity in the music. In this case, these loadings showed that music from the negative extreme tended to have a more complex tonal structure compared to music from the positive extreme.

Lastly, cross-validated classification showed that roughly 80% of the data points in this analysis were correctly classified, which indicated the accuracy of the discriminant model.

IV. DISCUSSION

The main objective of this paper was to explore how audio features can discriminate between different music preferences related to personality. This objective was accomplished via a series of complex steps that involved: 1) using Principle Components Analysis (PCA) to group participants' preference scores for specific music clips; 2) carrying out regression analyses to relate these groups derived from the PCA to various personality traits; and 3) conducting a Multiple Discriminant Analysis (MDA) to discriminate among these music preference groups using extracted audio features. While proceeding through these steps, hypotheses tested relations between music preferences and gender, age, music training, and familiarity with music clips. Results from the series of steps, stated as aims in the Introduction, provided a rich set of relations between music preferences, personality, and extracted audio features. These results will be discussed together in one section due to their inter-related nature. Before this section is given, however, results for the hypotheses mentioned above are discussed.

A. Music Preferences, Familiarity, and Demographics

Three hypotheses were stated in the Introduction, which were expressed to test expected relations between music preferences and: 1) familiarity, 2) gender and age, 3) and music training. Each of the tests done to investigate these relations provided significant results. Discussion about these results is presented in the order that the three hypotheses are given above.

As stated in the Introduction, the first hypothesis asserted that participants' familiarity is positively related to their preference score toward the same song. It was not surprising to find that this was the case. This relation proved to be quite strong (r = .62), which further emphasizes that people tend to seek out and play music that they like. It was good to formally test this relation since much of the previous research on music preferences and personality had not used music as auditory stimuli to measure music preferences (e.g., Delsing et al., 2008; George et al., 2007; Rentfrow & Gosling, 2003; Zweigenhaft, 2008), and those who had used music had not reported this relation (e.g., Cattell & Saunders, 1954; McCown et al., 1997). Nevertheless, it was important to maintain that familiarity with a given music clip is a necessary part of their preference for the same clip. For this reason, familiarity was not statistically controlled for when investigating the remaining results.

The second hypothesis was based on previous research that indicated music preferences were related to gender and age (e.g., Arnett, 1992; George et al., 2007; McCown et al., 1997; McNamara & Ballard, 1999; Zweigenhaft, 2008). With respect to gender, results agree with the majority of this research, indicating that music preferences are related to gender. Nonetheless, the present finding does not answer why this relation exists. Are there inherent aspects contained within certain music that makes it more appealing to men or to women? Or is this relation better explained by social norms that dictate what music men or women should find more appealing?

A similar conundrum can be raised about the significant interaction effect found between music preferences and age. Again, this finding does confirm previous research showing that music preference and age are related (e.g., Zweigenhaft, 2008). So, is this relation due to a generational cohort effect, or do music preferences change over time? Considering additional literature that has dealt with this issue (e.g., Delsing et al., 2008; Levitin, 2006), it is argued that the significant interaction effect found related to age is more likely due to a cohort effect.

Despite the relatively small effect sizes for both gender and age (partial η^2 = .03, and partial η^2 = .06, respectively), these variables were statistically controlled for in further analyses. This was done to ensure that there was a sufficiently accurate measure of music preferences regardless of gender or age. In addition, the specific nature of these relations was not provided because of space restrictions and of the peripheral nature of these results toward this paper's objective and aims.

The third and final hypothesis was posed in order to examine the relation between music preference and music training. This relation had received only limited attention previously (e.g., George et al., 2007), but the hypothesis was based on literature stating that music training has an effect on music perception (e.g., Lamont & Dibben, 2001; Levitin, 2006; McAdams & Matzkin, 2001). This statement was extended to suggest that music preferences might also be related to music training. This hypothesis was found to be true, which supports findings by George et al., but again, the effect size found for this interaction effect was small (partial η^2 = .04). Similar to the results for age and gender, however, music training was statistically controlled for in further analyses to help ensure a sufficiently accurate measure of music preferences. Furthermore, the specific nature of this relation was not provided because the peripheral nature of these results toward this paper's objective and aims.

B. Music Preferences, Audio Features, and Personality

There were three objectives specified for this paper, which represent three separate aspects of our main objective:

1. Develop a preliminary model of music preferences using audio stimuli (i.e., specific songs).
2. Relate these music preferences to personality traits.
3. Discriminate between music preferences using extracted audio features.

Due to their exploratory nature, there were no hypotheses directly connected to any of these aims. Nonetheless, there are rich results associated with each of these aims, which will be discussed here in the order that these aims have been specified. Furthermore, the results from these aims are very inter-related in nature. So, discussion concerning the overall interpretation of these results is provided at the end of this section.

For the first aim, Principle Components Analysis (PCA) was done to develop a preliminary model of music preferences. Ultimately, this PCA resulted in a 9-component model that accounted for a good proportion of the variance (61%) in participants' preference scores toward the music clips. What is particularly interesting about this model is how the music has been grouped among these components. For the most part, the music provided in Table 1 is grouped similar to how we would expect it to be grouped according to genre (c.f., Amazon.com, 2007; AMG, 2007). The abridged list of music presented in Table 1 is also representative of the remaining music used in this study. This would seem to support previous research that has used genre labels to measure music preferences (e.g., Delsing et al., 2008; George et al., 2007; Rentfrow & Gosling, 2003; Zweigenhaft, 2008). So, genre labels do seem provide a reasonable level of accuracy. Still, there are three reasons to suggest that greater accuracy can still be sought after. First, it should be kept in mind that music selected for this study was based on converging information from three different music industry sources (i.e., Amazon.com, 2007; AMG, 2007; Last.fm, 2007). This meant that the music selected for this study was likely more prototypical for each of the identified genres, rather than fringe music for these genres.

Second, despite the likelihood that most of the music contained in this study was more prototypical, there were instances where preference for specific music clips would bleed across music preference components. For example, the song Nookie, by Limp Bizkit was most often liked by participants who also reported enjoying music that would typically be classified as Hard Rock or Heavy Metal. Nonetheless, this particular song was also often liked by participants who reported enjoying other music that would typically be classified as Rap or Hip-Hop. There may have been instances when same participants liked both Heavy Metal and Rap music, but there was a sufficient number of times where this was not the case. Otherwise, Heavy Metal and Rap music would have fallen under the same component in the PCA, which is clearly not the case. There is even one instance shown in Table 1 where the music clip does not group with any of the music that it is categorized with according to genre. Vanessa-Mae is typically considered as a Classical artist (e.g., AMG, 2007), but here her song, Destiny, was more likely to be preferred by participants who also reported enjoying music that would be either classified as Pop or, alternatively, Dance.

Third, Pachet and Cazaly argue that there are consistency issues with respect to genre taxonomies, and in particular, with respect to the distinction between Rock and Pop music (as cited in Aucouturier & Pachet, 2003). Nonetheless, from Table 1, it appears that audiences might be able to adequately distinguish between several categories of Rock and Pop music, based on their music preferences. Therefore, greater accuracy could be gained by grouping music according to preference and not according to genre. This also offers an additional reward. By leveraging music preference groups and relating these groups to audio features, it might be possible to identify the relevant music groups according to how these groups are preferred by various audiences, and not by potentially arbitrary music classification according to various industry sources.

For the second aim, nine separate stepwise regressions were done to ascertain the relation between music preferences and personality. If part of the main objective was to improve our understanding of the relation between music preferences and personality, it was believed that relating music preferences to more detailed personality facets from the Big Five dimensions would help achieve this. Only one previous study is known to have looked at the relation between music preferences and personality at this more detailed level (i.e., Zweigenhaft, 2008). Nonetheless, over 200 correlations had been performed in that analysis, which makes it difficult to discern true significant findings from spurious ones. Though the analysis from the current study was exploratory in nature, the stepwise regression would help prevent spurious findings. With exception to results for component 6, re-tests of the results using a randomly drawn sub-sample indicated that these results were reliable. The R^2 values indicate medium effect sizes for each of the regression equations. Furthermore, there are comparable results to prior research that has investigated the relation between music preferences and personality. For instance, individuals who prefer Blues, Classical, or Jazz music tend to be more *Open to Experience* (O; Delsing et al., 2008; George et al., 2007; Rentfrow & Gosling, 2003), and specifically with respect to *Aesthetics* (Zweigenhaft). Also similar to Zweigenhaft was that participants who scored high in *Excitement-Seeking* tended to report enjoying Rap music. This latter finding could be due to the predominantly greater intensity of bass sounds in Rap (McCown et al., 1997), which will be further discussed shortly.

For the third aim, a Multiple Discriminant Analysis (MDA) used extracted audio features to discriminate music clips according to the nine music preference components. The first four discriminant functions accounted for the lion's share in the variability among the audio feature data points. These functions also seemed to give better insight into how audio features could express music preferences more accurately, and how these features could be linked to personality. Thus, the following discussion considers these MDA results in light of the other results concerning music preferences and personality. In addition, there are many possible examples that highlight how audio features could express music preferences and its link to personality. For brevity, this discussion provides two examples.

The first function from the MDA provides a clear distinction between Rap music on the one hand, and Classical music on the other hand. Given the function loadings, it was clear that one of the features used by this function was related to the amount of percussive sounds in the audio. Comparing the MDA results to the results relating music preferences with personality, it seems that *Excitement-Seeking* tends to be negatively related to scores along this function. That is, lower function scores tended to come from music like Rap, which tended to be enjoyed by participants' who were higher in *Excitement-Seeking*. The reverse was true of higher function scores, which tended to come from music like Classical. This helps further explain the results from McCown et al. (1997).

The third function provides a clear distinction between Jazz music on the one hand, and Pop music on the other hand. Given the function loadings, it was clear that tonal complexity was a key feature used by this function to discriminate these groups according to extracted audio features. Further comparisons between the MDA results and results relating music preferences with personality indicated that *Aesthetics* seems important in this distinction. In this case, lower function scores tended to come from music like Jazz, which tended to be enjoyed by participants' who were higher in their openness to aesthetic experiences, like attending the ballet or an art show. Higher function scores tended to come from music like Pop, which tends to have a simpler tonal complexity and is enjoyed more by participants who were lower in *Aesthetics*.

These results provide some good preliminary examples that result from considering the role that audio features play in the relation between music preferences and personality. As part of the conclusion it is argued that this research demonstrates how music preferences can be measured more accurately, and so, helps provide greater insight into the relation between music preferences and personality.

C. Limitations and Future Directions

The single largest limitation with the present study is linked to its exploratory nature. The models derived and presented here need to be tested further for reliability and validity. Furthermore, the role of cultural background or nationality was not investigated here, primarily due to the large number of different nationalities represented in the sample after accounting for participants from the U.S., Canada, and the U.K. Both these limitations could be addressed by conducting a study with the same audio stimuli using a new sample of participants from one specific country or area of the world.

Results regarding audio extracted features have also been presented in this study, but the specific nature of these features remains undisclosed due to proprietary reasons. To verify these described results, further studies investigating the relations between music preferences, audio features, and personality should be done using open source software, such as the MIR toolbox (Lartillot & Toiviainen, 2007).

V. CONCLUSION

The current paper has addressed the genre ambiguity problem by exploring how extracted audio features can be used to distinguish among preferences to specific music clips. Results show that while genre does provide a reasonable level of accuracy to measure music preferences, there are cases where songs are often enjoyed by audiences of two different genres. The initial models revealed distinct personality traits that are related to certain music preferences, which can be discriminated by using extracted audio features. Therefore, these results reveal that audio features can provide greater accuracy for predicting music preferences.

In a similar vein, the technique demonstrated in this paper provides the opportunity to achieve still greater insight into what audio features could be specifically preferred by people with certain personality traits. In turn, this could improve our understanding of what music we like and why. From an applied perspective, this might also be used to improve various technologies, such as recommender systems or intelligent awareness systems.

ACKNOWLEDGMENT

We thank Marie Curie Actions for funding this research as part of the CONTACT project (FP6-008201). We'd also like to thank Janto Skowronek, Don Bouwhuis, Boris de Ruyter, William Green, Marco Tiemann, Alberto Novello, among others for lending ideas, comments, and generally supporting this research. A special thanks to our participants for dedicating the time to complete this experiment.

REFERENCES

Amazon.com. (2007). Retrieved October 10, 2007, from http://www.amazon.com

AMG. (2007). *All music guide*. Retrieved October 10, 2007, from http://www.allmusic.com

Arnett, J. (1992). The soundtrack of recklessness: Musical preferences and reckless behavior among adolescents. *Journal of Adolescent Research, 7*(3), 313-331.

Aucouturier, J.-J., & Pachet, F. (2003). Representing musical genre: A state of the art. *Journal of New Music Research, 32*(1), 83-93.

Cattell, R. B., & Anderson, J. C. (1953). The measurement of personality and behavior disorders by the I.P.A.T. music preferences test. *The Journal of Applied Psychology, 37*(6), 446-454.

Cattell, R. B., & Saunders, D. R. (1954). Musical preferences and personality diagnosis: I. A factorization of one hundred and twenty themes. *The Journal of Social Psychology, 39*, 3-24.

Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.

Delsing, M. J. M. H., Ter Bogt, T. F. M., Engels, R. C. M. E., & Meeus, W. H. J. (2008). Adolescents' music preferences and personality characteristics. *European Journal of Personality, 22*(2), 109-130.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286-299.

George, D., Stickle, K., Rachid, F., & Wopnford, A. (2007). The association between types of music enjoyed and cognitive, behavioral, and personality factors of those who listen. *Psychomusicology, 19*(2), 32-56.

John, O. P., & Srivastava, S. (1999). The big five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102-138). New York: Guilford Press.

Juslin, P. N. (2000). Cue utilization in communication of emotion in music performance: Relating performance to perception. *Journal of Experimental Psychology: Human Perception and Performance, 26*(6), 1797-1813.

Juslin, P. N., & Sloboda, J. A. (Eds.). (2008). *Music and emotion: Theory and research*. New York, NY: Oxford.

Lamont, A., & Dibben, N. (2001). Motivic structure and the perception of similarity. *Music Perception, 18*(3), 245-274.

Lartillot, O., & Toiviainen, P. (2007). A matlab toolbox for musical feature extraction from audio. *Proceedings of the 10th International Conference on Digital Audio Effects* (*DAFx-07*, pp. 237-244). Bordeaux, France.

Last.fm. (2007). Retrieved October 10, 2007, from http://www.last.fm

Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. New York: Dutton/Penguin.

Litle, P., & Zuckerman, M. (1986). Sensation seeking and music preferences. *Personality and Individual Differences, 7*(4), 575-577.

McAdams, S., & Matzkin, D. (2001). Similarity and invariance and musical variation. *Annals of the New York Academy of Sciences, 930*, 62-76.

McCown, W., Keiser, R., Mulhearn, S., & Williamson, D. (1997). The role of personality and gender in preference for exaggerated bass in music. *Personality and Individual Differences, 23*(4), 543-547.

McNamara, L., & Ballard, M. E. (1999). Resting arousal, sensation seeking, and music preference. *Genetic, Social, and General Psychology Monographs, 125*(3), 229-250.

Negus, K. (1996). *Popular music in theory*. Oxford, England: Blackwell Publishing.

North, A. C., & Hargreaves, D. J. (1999). Music and adolescent identity. *Music Education Research, 1*(1), 75-92.

Nunnally, J. (1967). *Psychometric theory*. New York, NY: McGraw-Hill.

Pampalk, E. (2004). A matlab toolbox to compute similarity from audio. *Proceedings of the International Conference on Music Information Retrieval* (*ISMIR '04*, pp. 254-257). Barcelona, Spain.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of Personality and Social Psychology, 84*(6), 1236-1256.

Rentfrow, P. J., & Gosling, S. D. (2006). Message in a ballad: The role of music preferences in interpersonal perception. *Psychological Science, 17*(3), 236-242.

Rusting, C. L. (1998). Personality, mood, and cognitive processing of emotional information: Three conceptual frameworks. *Psychological Bulletin, 124*(2), 165-196.

Schwartz, K. D., & Fouts, G. T. (2003). Music preferences, personality style, and developmental issues of adolescents. *Journal of Youth and Adolescence, 32*(3), 205-213.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn and Bacon.

Zweigenhaft, R. L. (2008). A do re mi encore: A closer look at the personality correlates of music preferences. *Journal of Individual Differences, 29*(1), 45-55.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*(3), 432-442.