A CROSS-CULTURAL STUDY OF MUSIC MOOD PERCEPTION BETWEEN AMERICAN AND CHINESE LISTENERS

Xiao Hu

Faculty of Education
The University of Hong Kong
xiaoxhu@hku.hk

Jin Ha Lee

The Information School
University of Washington
jinhalee@uw.edu

ABSTRACT

Music mood has been recognized as an important access point for music and many online music services support browsing by mood. However, how people judge music mood has not been well studied in the Music Information Retrieval (MIR) domain. In particular, people's cultural background is often assumed to be an important factor in music mood perception, but this assumption has not been verified by empirical studies. This paper reports on a study comparing mood judgments on a set of 30 songs by American and Chinese people. Results show that mood judgments do indeed differ between American and Chinese respondents. Furthermore, respondents' mood judgments tended to agree more with other respondents from the same culture than those from the other group. Both the song characteristics (e.g., genre, lyrical or instrumental) and the non-cultural background of the respondents (e.g., age, gender, familiarity with the songs) were analyzed to further examine the difference in mood judgments. Findings of this study help further our understanding on how cultural background affects mood perception. Also discussed in this paper are implications for designing MIR systems for cross-cultural music mood classification and recommendation.

1. INTRODUCTION

The number of studies on music mood has been increasing in the Music Information Retrieval (MIR) domain as many perceive music mood as a potential feature for organizing and recommending music. However, previous research asking people to provide mood tags for short music clips found that it is a highly subjective feature and the vocabulary of music mood varies widely among users [8]. In addition to the features inherent in music itself there are a number of features that can affect how people determine the mood of music (e.g., their current state of mind, life events). We believe that one important factor is the cultural context of the user. However, currently there are no cross-cultural studies that specifically compare how people perceive and determine the mood of music in

the MIR domain.

In this study, we explore if users from China and the United States perceive music mood in different ways. We chose to compare these cultures for several reasons. First, China is a dominant Eastern culture while the United States is a dominant Western culture. Second, although the influence of American pop culture is gradually increasing in China, due to historical and political factors, Chinese people are far less affected by Western culture as compared to people from other East Asian countries such as Korea or Japan [13]. Third, one of the authors is fluent in Chinese and English, which is important as translating the mood labels while preserving the subtle nuances can be challenging for non-native speakers.

2. LITERATURE REVIEW

2.1 Cross-Cultural Studies in Music Psychology

There are a number of studies in music psychology on comparing responses on Western and non-Western music from Western and non-Western listeners but many of them focused on aspects such as memorability of music (e.g., [9]) or perception of complexity (e.g., [2]) rather than perception of music moods.

Balkwillm and Thompson [1] are among the first investigating whether judgments on music mood can transcend cultural boundaries. They recruited 30 Western listeners to judge 12 Hindustani raga excerpts in order to see if people can identify the intended emotion in music from an unfamiliar tonal system. Their findings showed that the emotions of joy, sadness and anger (but not peace) were identifiable by the listeners and the emotion judgments were significantly related to psychophysical characteristics of the pieces. However, their study did not compare judgments of people from different cultures.

Gregory and Varney [5] compared mood descriptors applied to Indian classical, Western classical, and new age music by Indian and European listeners and revealed "many subtle differences in affective response" in mood descriptors. They suggested that "the affective response to music is determined more by cultural tradition than by the inherent qualities of the music (p.47)." More recently, Wong et al. [14] found that Indian and Western listeners showed in-culture bias when judging the tension in Western and Indian music. However, Fritz et al. [4] found that native African (Mafa) listeners could recognize three basic emotions (happy, sad, scared/fearful) expressed in

Western music with above chance accuracy and suggested that, "the expression of these basic emotions in Western music can be recognized universally (pp.253)." The conflicting results in these studies highlight the need for more empirical research on cross-cultural music mood perception. In addition, these studies generally focused on classical and/or ethnic music whereas our study focuses on popular music.

2.2 Cross-Cultural Studies in Music Information Retrieval

In the MIR domain, there are few studies that examined cross-cultural aspects related to how users interact with and search for music. Lee et al. [7] collected music related questions from Q&A websites based in North America (i.e., Google Answers) and Korea (Naver Knowledge-iN) and did a comparative analysis. They found that Korean users experienced a number of challenges in cross-cultural/multilingual music searches: 1) they often failed to provide bibliographic metadata such as composer, performer or title in their queries, 2) they had difficulties in using Western music genres and instead relied on association-based concepts (i.e. where the music was used), and, 3) they had difficulties in using and transliterating lyrics information. The authors suggest that new access points for accommodating cross-cultural/multilingual music searching including associate metadata (i.e., usage) are necessary.

Nettamo et al. [11] also conducted a cross-cultural study of mobile music retrieval, management, and consumption behaviors of people in New York vs. Hong Kong. They found several differences in how music was being managed, shared, and used. For example, New Yorkers sought music information through various channels including blogs, websites, magazines, etc., and mood and context of use were factors affecting how they generated playlists. Hong Kong users, on the other hand, sought for music only through limited channels such as ranking websites or through friends, and did not use playlists at all.

Both of these studies indicate that there were in fact differences between users from different cultures with regards to their music related behaviors. However, none of these studies examined how they perceive music mood which is the gap this work is attempting to bridge.

3. STUDY DESIGN

This study focuses on the following set of research questions:

1. Do Americans and Chinese have different perceptions of mood on the same set of popular songs? In other words, do people from the same culture tend to agree more with each other?
2. Do some moods tend to be more agreeable among people from certain culture?
3. Do characteristics of the songs (i.e., genre, vocal or instrumental) affect the difference on mood judgments between Americans and Chinese?

4. Do users' non-cultural background (i.e., gender, age, and familiarity with the songs) affect the difference on mood judgments between Americans and Chinese?

To answer these questions, we created an online survey in which each user listened to thirty 30-second music clips and selected the most appropriate mood cluster among the five given clusters for each piece. We used the same mood clusters that are used in MIREX [1] Audio Mood Classification (AMC) Task [6] (reprinted in Table 1). Users can choose "other" if they think none of the mood clusters is applicable to the music piece.

Cluster 1 (C_1)	passionate, rousing, confident, boisterous, rowdy
Cluster 2 (C_2)	rollicking, cheerful, fun, sweet, amiable/ good natured
Cluster 3 (C_3)	literate, poignant, wistful, bittersweet, autumnal, brooding
Cluster 4 (C_4)	humorous, silly, campy, quirky, whimsical, witty, wry
Cluster 5 (C_5)	aggressive, fiery, tense/anxious, intense, volatile, visceral

Table 1. Five Mood Clusters used in MIREX [6].

Half of the 30 pieces were selected from the MIREX AMC task test collection with the help of IMIRSEL. This AMC test collection was created based on the APM (Associated Production Music)[2] collection, and covers a variety of different music genres. The mood of each piece in the AMC test collection was judged by three MIREX evaluators whose cultural backgrounds ranged from Europe, America and Asia [6]. In order to avoid including the "obvious" examples that received high agreement from evaluators in the test data set, we selected the songs for which there was greater disagreement among the MIREX evaluators. As these pieces were instrumental without a vocal part, we balanced our test set by drawing the other half of the 30 pieces from the USPOP collection [3] and ensured that they all had vocal components. The songs from the USPOP collection were chosen similarly to the APM songs; we selected songs that had greater disagreement in the mood judgments from six IMIRSEL members (c.f., [6]).

We recruited two user groups for the survey: people who were raised in Mainland China and considered themselves "Chinese", and people who were raised in the United States, and considered themselves "American". All subjects were recruited from large universities in the U.S. The survey was deployed in both Chinese (Mandarin) and English; all the mood labels in English were translated into Chinese by the first author for the Chinese survey. In both surveys, we asked the users if they had heard the

---

music clip before and if they could name the artist and the song title in order to gauge their familiarity with the songs. Figure 1 shows the screenshot of the online survey in English.

Figure 1. Survey interface for American listeners

4. DATA AND DISCUSSION

4.1 Overview

There were a total of 55 responses from Chinese and 45 from Americans; however, not all responses were complete. 31 listeners completed the Chinese and American surveys, respectively, for a total of 62 complete responses. Table 2 shows the demographic information of the selected respondents. Among the Chinese respondents; 23 of them had been living in the U.S. for less than 2 years, 3 had been in the U.S. for 3-5 years, 1 for 6-8 years, and 4 for 9-11 years.

Cultural background	Age			Gender	
	Min	Max	Avg.	Male	Female
American	22	55	31.8	6	25
Chinese	19	46	26.2	10	21

Table 2. Demographics of survey respondents

In the following data analysis, chi square (χ^2) statistics are used to test whether distributions of two categorical variables (e.g., cultural background, mood categories) are independent from each other [12] unless noted otherwise. In the following subsections, we will answer each research question based on analysis of the survey responses.

4.2 Difference Between Cultural Groups

4.2.1 Mood Judgments on All Songs

Figure 2 shows the distribution of mood judgments of the two groups on all 30 songs. Chinese users selected mood Cluster 1 (*passionate, rousing, confident, boisterous,*

rowdy) and Cluster 3 (*literate, poignant, wistful, bitter-sweet, autumnal, brooding*) more often than Americans whereas Americans chose Cluster 2 (*rollicking, cheerful, fun, sweet, amiable/good natured*) more often than Chinese. Both groups had similar numbers of judgments on Cluster 4 (*humorous, silly, campy, quirky, whimsical, witty, wry*) and Cluster 5 (*aggressive, fiery, tense/anxious, intense, volatile, visceral*). More Americans than Chinese chose the "other" option. A chi square test indicates that listeners' selection of mood clusters significantly depends on the cultural group they belong to ($\chi^2 = 73.64$, df = 5, $p < 0.0001$). In other words, there was a significant difference between Americans and Chinese when examining their mood judgments as a whole. A follow-up Tukey multiple comparison test [15] showed that judgments of the two cultural groups on the "other" cluster were significantly different from those on all the other clusters (at $p < 0.05$), judgments on Cluster 2 were significantly different from those on Cluster 1 and 3 (at $p < 0.05$), and there were no significant differences between judgments on other pairs of mood clusters.
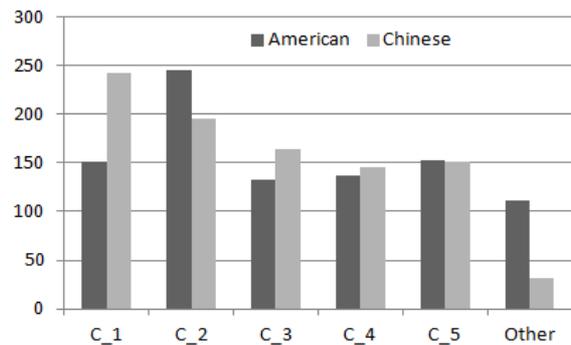
Figure 2. Distribution of mood judgments of the two groups

4.2.2 Agreement on Mood Judgments

In order to find out whether listeners from the same cultural group would agree more with each other than with listeners from another cultural group, we calculated the level of agreement on mood judgments among individual listeners. For categorical data such as the mood judgments, agreement is typically calculated based on the Sokal-Michener coefficient, which is a ratio of the number of pairs with the same values and the total number of variables (songs in this case) [12]. For instance, if two listeners *i* and *j* had the same mood judgment on 15 of the 30 songs, the agreement ratio between them will be 0.5. Table 3 shows the average agreement ratio among pairs of listeners within and across cultural groups in this study. Within each cultural group, users show 0.35 agreement rate. However, across cultural groups the agreement rate drops to 0.30. A non-pair wise *t*-test was conducted to test the significance of the difference on agreement ratio within each cultural group and across cultural groups. Both tests revealed a statistically significant difference. Therefore, our data support the hypothesis that listeners

tended to agree more with others from the same cultural background than those from another cultural background.

	American	Chinese	T statistics	p-value
American	0.35	0.30	11.44	<0.001
Chinese	0.30	0.35	12.24	<0.001

Table 3. Average agreement ratio within and between cultural groups

4.3 Mood Clusters vs. Cultural Groups

We also investigated which mood clusters received higher agreement from people in each cultural group. We examined all pairs of responses from each group. Since there were 31 responses from each group, there were a total of 465 pairs of responses within each group. Each response had 30 mood judgments, thus there were 13,950 pairs of judgments in each group. Between the two cultural groups, there are 31 * 31 = 961 pairs of responses and 961 * 30 = 28,830 pairs of judgments. Table 4 lists the number of agreed pairs of judgments on each mood cluster within each cultural group and across cultural groups. It shows that American listeners agreed more on Cluster 2 and 5 while Chinese listeners agreed more on Cluster 1 and 3. The difference between the two groups is statistically significant ($\chi^2 = 668$, df = 5, $p < 0.0001$).

	C_1	C_2	C_3	C_4	C_5	Other	Total
American	706	1477	778	587	1094	270	4912
Chinese	1355	995	1203	443	894	11	4901
Across	1704	2122	1713	881	1999	131	8550

Table 4. Number of agreed pairs of judgments across mood clusters

4.4 Song Characteristics vs. Cultural Groups

Half of the test songs were instrumental and the other half were vocal. The two cultural groups showed significant difference in judging the mood for both instrumental ($\chi^2 = 88.09$, df = 5, $p < 0.0001$) and vocal songs ($\chi^2 = 28.98$, df = 5, $p < 0.0001$). Table 5 shows the agreement ratios among all judgment pairs on instrumental and vocal songs. Lyrics definitely seem to help achieve a higher agreement for Americans while they have essentially no effect on Chinese (it should be noted that all lyrics were in English). In addition, cross-culturally the two groups were more likely to provide different judgments on instrumental pieces than vocal ones. As discussed in [8], even if Chinese listeners cannot comprehend the lyrics as well as American listeners, the delivery of the singer may still affect how they determine the mood of the song.

	Instrumental	Vocal	All
American	0.28	0.41	0.35
Chinese	0.36	0.35	0.35
Across	0.25	0.34	0.30

Table 5. Agreement ratio on instrumental vs. vocal songs

We also looked at the genres of the songs as provided by APM and USPOP. Table 6 shows the genre distribution of the songs as well as instrumental vs. vocal information. Dance and Easy-listening songs were all instrumental while songs in the remaining genres were mostly vocal. For each of the five genres, mood judgments were significantly dependent on cultural groups ($\chi^2 = 21.91 \sim 46.68$, df = 5, $p < 0.001$). Table 6 also shows the agreement ratios across genres. As it can be seen, Americans agreed more on Pop songs whereas Chinese agreed more on songs in Other and Easy-listening. Cross-cultural agreement levels are generally lower than those within cultural groups. Among all the genres, Dance and Easy-listening songs had the least cross-cultural agreement.

	Dance	Easy-listening	Pop	Rock	Other	Total
Instru.	4	5	2	1	3	15
Vocal	0	0	5	7	3	15
American	0.30	0.29	0.46	0.35	0.31	0.35
Chinese	0.29	0.38	0.32	0.35	0.41	0.35
Across	0.22	0.28	0.33	0.31	0.30	0.30

Table 6. Song distribution and agreement ratio across genres

4.5 Listener Characteristics vs. Cultural Groups

The aggregated mood judgments across songs and cultural groups were still statistically significant when we consider the gender of the listeners (i.e., Chinese male vs. US male, $\chi^2 = 18.28$, df = 5, $p = 0.0026$; Chinese female vs. US female, $\chi^2 = 52.83$, df = 5, $p < 0.0001$).

As Table 2 shows, the Chinese respondents in this study were generally younger than the American respondents. To minimize the possible influence of age on mood judgments, we compared the answers from listeners of the same age range (22-46 years old) in both culture groups (24 Chinese and 28 American). The mood judgments of two cultural groups were still significantly different ($\chi^2 = 61.85$, df = 5, p < 0.0001).

In this study, a listener's familiarity with a song is measured by their answers to two questions: 1) whether he or she had heard the song before; and 2) whether he or she can identify the artist name and song title. A "no" answer to both questions indicates low familiarity, a "yes" to both questions indicates high familiarity, and a "yes" and a "no" indicates medium familiarity. The reason for using these two questions instead of directly asking the listeners their level of familiarity is because people may have different interpretations on song familiarity. Some people might consider a song familiar if it invokes any memory while other people might not think it is familiar unless they could actually sing part of the song. The two questions are objective, and thus are easier to answer and avoid personal biases. Table 7 shows the distribution of the level of familiarity across Americans and Chinese. As the test songs were Western songs, it is not surprising that American listeners were more familiar with the songs than Chinese listeners.

In order to see whether the level of familiarity has an effect on mood judgment agreement, we calculated the agreement ratio with various combinations of familiarity

levels in each cultural group as well as across cultural groups. Each cell in Table 8 shows the agreement ratio among all judgment pairs with corresponding familiarity levels.

	Unfamiliar	Medium	Familiar	N/A	Total
American	617	120	192	1	930
Chinese	836	75	14	5	930

Table 7. Distribution of the level of familiarity

		Unfamiliar	Medium	Familiar
American	Unfamiliar	0.32	0.36	0.36
	Medium	-	0.43	0.44
	Familiar	-	-	0.44
Chinese	Unfamiliar	0.35	0.37	0.23
	Medium	-	0.36	0.22
	Familiar	-	-	0.24
Across	Unfamiliar	0.28	0.31	0.29
	Medium	-	0.37	0.33
	Familiar	-	-	0.24

Table 8. Agreement ratio across different levels of familiarity

For American listeners, being familiar with the songs did improve the odds of agreeing. However, medium and high familiarity did not appear to have much effect on agreement. For Chinese listeners, having heard the songs before (medium familiarity) slightly increased the agreement ratio, but high familiarity with the songs appears to actually decrease agreement, which is unintuitive. We suspect that this might be due to the sparseness of the samples: there were only 14 (out of 930) cases where the Chinese listeners were highly familiar with a song [Table 7]. It may also indicate that a mere identification of title and/or artist name from Chinese listeners does not imply that they understood what the song was about. Table 8 also shows that the agreement ratios between Americans and Chinese are lower than those of Americans.

5. DISCUSSION

Our analysis suggests that there is in fact a significant difference between how Americans and Chinese perceived music mood. From the total number of mood judgments across mood clusters, Chinese listeners chose Cluster 1 more often than American listeners. We conjecture that this difference may be attributed to the differences between Chinese and Western cultures. In Chinese culture, people tend to restrain the expression of feelings and Chinese people are generally more introverted compared to Western people [9], and thus may be more likely to think a Western music piece is "passionate," "rousing," or "boisterous" (Cluster 1). Previous research also found that Chinese value low-arousal positive affect (e.g., calm) whereas Westerners value high-arousal positive affect [11]. This may help explain the higher responses on Cluster 3 for Chinese listeners and Cluster 2 for American listeners. When the mood of the song is not clear, people

may end up selecting moods that they generally prefer since they are more likely to focus on those moods; in other words, they hear what they want to hear.

It is also interesting to see fewer judgments of "other" among Chinese users. This may be related to the collectivism commonly seen in Eastern cultures and the individualism in Western cultures. Chinese listeners used one of the given five mood clusters 96.6% of the time. However, American listeners disagreed with the presented mood clusters more often, using the "other" option for 11.9% of their judgments, more than 3 times as often as Chinese.

These findings have implications for designing MIR systems for people with different cultural backgrounds. For example, a mood classification system may be designed so that it treats songs with mixed moods in a different way: categorizing them into Cluster 2 for Americans vs. Cluster 1 or 3 for Chinese reflecting their expectations. In addition, while we generally think it is more user-friendly to let users browse for music with different moods than asking users to search with their own mood terms, this would appear to be even more critical for Chinese listeners, as they seem to prefer using given organizational structures rather than providing their own input (via the "other" option as shown in Figure 1).

Among all the 30 test songs, the one with the highest disagreement between Americans and Chinese was *Got to get you into my life* by The Beatles. Figure 3 shows the judgment distribution across mood clusters for this song. Most of the Americans answered Cluster 2 while the Chinese' answers were spread out across multiple mood clusters with most answers in Cluster 4. A closer look at the data revealed that none of the Chinese listeners had listened to this song prior to this survey whereas 29 out of 31 American listeners had listened to this song and 19 of them were familiar enough to this song that they could name the artist. Cultural background evidently played an important role in mood judgments on this song. The Beatles are a symbol of the Western pop culture. Western listeners had probably been influenced by their background knowledge of the song and the band that this song *should* express a rollicking and cheerful (Cluster 2) or passionate (Cluster 1) mood. It is also possible that they were able to provide a mood judgment based on the whole song rather than the 30 seconds clip provided in the survey. In contrast, the Chinese listeners had no prior influence on how *others* had felt about this song or rest of the Beatles' songs, thus their answers were not as consistent as those from the American listeners.

Lyrics seem to affect how the Americans judge the mood of the songs, but not the Chinese [Table 5]. Although most of the Chinese listeners in this study could understand English, the music pieces in the survey are probably too short (30 seconds) for them to fully comprehend the lyrics and use them in mood judgments. Genre also affects people's agreement on music moods, but in different ways for different cultural groups. The fact that Americans agreed more on Pop songs is possibly related to their being familiar with the songs.
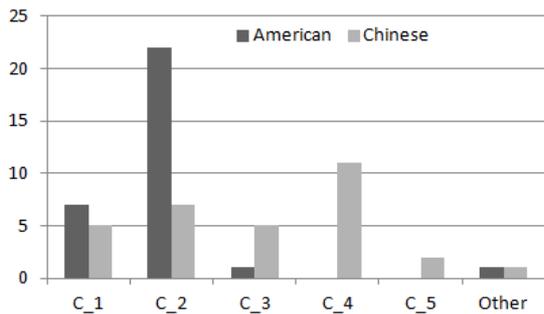
Figure 3. Mood judgment distributions for *Got to get you into my life*

As previously discussed, listeners' familiarity with the songs affects the level of agreement on mood judgments, but the influence is much stronger within the same cultural group. This makes it challenging to build an MIR system for users with cultural backgrounds that are different from the particular culture the music is from. For cross-cultural MIR systems, perhaps more flexibility should be provided to users. It may help to allow users to provide annotations so that they can complement the given "correct" mood labels. In such systems, it will be possible for users to assign multiple mood labels to a song, change a song's mood labels, or add alternative labels to the songs.

6. CONCLUSION AND FUTURE WORK

In this paper, we presented a study comparing mood judgments on a common set of Western music pieces by American and Chinese listeners. Listeners from the two cultural groups indeed have different mood judgments and they tended to agree more with users from the same cultural group. Some genres seemed to be more difficult to reach user agreement across two groups, although further studies with a larger music samples should be conducted to validate the result. The cultural difference persists even when we consider the age and gender. The listeners' familiarity with the songs had a positive influence on the agreement level among users from the same cultural background of the songs. Findings of this study not only help further our understanding on how cultural background affects mood perception, but also have implications for designing cross-cultural MIR systems.

It should be noted that the Chinese respondents in this study have lived in the U.S., but there were not enough data to analyze the influence of this factor on their music mood judgments. In our future study, we will collect responses from Chinese people in China and compare the results. We will also investigate why people assign tracks to certain mood clusters by conducting in-depth interviews. In addition, we plan to increase the diversity of our user group by including users from countries other than China and United States such as Korea. Although Korea also represents non-Western culture, and Chinese and Korean cultures historically share a great deal of similarities, Korea is much more heavily influenced by American pop culture than China. Thus, comparing user groups from these countries may provide insights into how the exposure to other pop culture can affect the way people perceive the mood of music.

7. REFERENCES

[1] L. Balkwillm and W. F. Thompson: "A Cross-cultural investigation of the perception of emotion in music: psychophysical and cultural cues," *Music Perception*, Vol. 17, No. 1, pp. 43-64, 1999.

[2] T. Eerola, T. Himberg, P. Toiviainen and J. Louhivuori: Perceived complexity of Western and African folk melodies by Western and African listeners, *Psychology of Music*, Vol. 34, No. 3, pp. 337-371, 2006.

[3] D. Ellis, A. Berenzweig and B. Whitman: "The USPOP 2002 Pop Music Data Set," Retrieved from http://labrosa.ee.columbia.edu/projects/musicsim/uspop2002.html, 2003.

[4] T. Fritz, S. Jentschke, N. Gosselin, D. Sammler, I. Peretz, R. Turner, A. D. Friederici and S. Koelsch: "Universal recognition of three basic emotions in music," *Current Biology,* Vol.19, pp.573–576, 2009.

[5] A. H. Gregory and N. Varney: "Cross-cultural comparisons in the affective response to music," *Psychology of Music*, Vol. 24, pp. 47-52, 1996.

[6] X. Hu, J. S. Downie, C. Laurier, M. Bay and A. F. Ehmann: "The 2007 MIREX Audio Mood Classification task: Lessons learned," *Proceedings of the 9th International Society for Music Information Retrieval (ISMIR) Conference*, pp. 462-467, 2008.

[7] J. H. Lee, J. S. Downie, and S. J. Cunningham: "Challenges in cross-cultural/multilingual music information seeking," *Proc. of ISMIR*, pp. 1-7, 2005.

[8] J. H. Lee, T. Hill, and L. Work: "What does music mood mean for real users?" *Proceedings of the iConference*, 2012.

[9] S. J. Morrison, S. M. Demorest and L. A. Stambaugh: "Enculturation Effects in Music Cognition: The Role of Age and Music Complexity," *Journal of Research in Music Education*, Vol. 56, No. 2 pp. 118-129, 2008.

[10] R. R. McCrae, P. T. Costa, Jr. and M. Yik: "Universal aspects of Chinese personality structure," In M. H. Bond (Ed.) *The handbook of Chinese psychology*. Hong Kong: Oxford University Press, pp.189-207, 1996.

[11] E. Nettamo, M. Norhamo, and J. Häkkilä: "A cross-cultural study of mobile music: Retrieval, management and consumption," *Proceedings of OzCHI 2006*, pp. 87-94, 2006.

[12] R. R. Sokal and C. D. Michener: "A statistical method for evaluating systematic relationships," *University of Kansas Science Bulletin,* Vol. 38, pp. 1409–1438, 1958.

[13] J. L. Tsai, B. Knutson and H. H. Fung: "Cultural variation in affect valuation," *Journal of Personality and Social Psychology*, Vol.90, No.2, pp. 288–307, 2006.

[14] P. C. M. Wong, A. K. Roy and E. H. Margulis: "Bimusicalism: The implicit dual enculturation of cognitive and affective systems," *Music Percept*, Vol. 27, No. 2, pp. 81–88, 2009.

[15] J. H. Zar: *Biostatistical Analysis*, Fourth Edition, Prentice Hall, 1999.