KANSEI EVALUATION OF JAPANESE MUSIC CONSIDERING ARTIST'S GENERATION AND LISTENER'S AGE

Hiroshi Takenouchi[1a], Masaki Kawakubo[a]

[a] *Fukuoka Institute of Technology, Japan, h-takenouchi@fit.ac.jp*

ABSTRACT

We analyze trends in music listener preferences using Kansei evaluation and taking into account differences in listener age and the generation of the artist. This study conducts Kansei evaluation experiments with participants ranging from their late teens to their forties, who listen to music performed by male dance performance groups considering of experienced and younger artists. We employ impression evaluations for various types of music using the semantic differential method, and we analyze the results through factor analysis and clustering. The results reveal that the factors that express impressions of music by experienced and younger groups differed from each other. In addition, we compared impressions between music-only presentations and music videos that added visual information to the audio, confirming that the impressions did not differ between the two conditions.

Keywords: Kansei evaluation, Semantic differential method, Japanese pops, Multivariate analysis.

1 INTRODUCTION

Musical trends (genres and styles) change over time. Such changes stem from differences in the ways that listeners of different age groups perceive music, as well as variations in expression influenced by generational differences among the artists who sing or perform it. Therefore, even when artists belong to the same category—such as dancer performers or idols—the impressions of their music can differ depending on the artist' and the listeners' generation.

Previous studies of the impression analysis of music have included the Kansei evaluation of key modulation in classical music (Hasegawa, 2025) and in background music (BGM) (Kampfer, 2010), the analysis of music appreciation in terms of affective evaluation and neural activity (Kakegawa, 2012), and Kansei evaluations of drum rhythms (Kurosaki, 2011). These studies focused on listeners' impressions and demonstrated how music is received: however, they did not analyze musical impressions in relation to generational differences among artists and listeners.

[1]Corresponding author

Considering these differences allow affective evaluations to be analyzed from multiple perspectives—such as how younger listeners perceive the music of experienced artists, or how middle-aged listeners respond to songs by younger performers.

This study conducts Kansei evaluations that consider listener age and artist generation to analyze trends in listeners preferences. We focus on listeners aged from their late teens to their forties. Careful consideration is required when selecting artist generations. For example, when evaluating songs by a younger jazz performer and an experienced rock musician, the fundamental differences in musical atmosphere and genre make direct comparison difficult. Therefore, it is necessary to select both experienced and younger artists whose music belongs to the same genre or category. In this study, we focus on male dance performance groups affiliated with love, dream, happiness (LDH) Japan Inc., which represents a wide range of artists, from emerging talents to seasoned veterans.

In this experiment, Kansei evaluation is conducted using the semantic differential (SD) method (Osgood, 1964), based on debut, most recent, and representative songs of six artist groups (three experienced and three younger groups) by listener participants from their late teens to their forties. The experiment analyzes trends in music appreciation between experienced and younger groups based on the Kansei evaluation results.

Through the experiment, we aim to clarify the following three points:

1. Whether the impression of a song changes depending on the listener's age.
2. Whether the impression of a song changes depending on the artist's generation.
3. Whether the impression of a song changes when visual elements are added to the audio.

Items 1 and 2 are the main focus of this study, involving a comparative investigation of impression evaluation tendencies based on differences in listener age and artist generation. Regarding item 3, since music videos (MVs), which are often released alongside songs, may significantly impact on listener impressions, this study examines differences in impressions when evaluating the audio-only versions and those from MVs that include visual information.

2 EVALUATION EXPERIMENTS

2.1 Outline

We conducted two Kansei evaluations: one to assess listeners' impressions when hearing the music alone, and another to evaluate their impressions when visual information was added in a MV. In Experiment 1, we presented only a 15-second excerpt from the chorus section of each artist's debut and latest songs and examined age-based differences in listener impressions through factor analysis and cluster analysis using the farthest neighbor method. In Experiment 2, we presented each artist's representative song in two formats—audio-only and audio with visual content (MV)—and investigate age-based differences in impressions using factor analysis.

The factor analysis parameters are as follows: the factor extraction method uses bartlett method, the rotation method of factor axis uses Varimax rotation, and the number of factors is

three in the Experiment 1 and two in the Experiment 2 decided by scree plot results of eigen values of the correlation matrix (Kaiser standard). Furthermore, as mentioned in Sections 2.2 and 2.3, the results of this study were analyzed using exploratory factor analysis because the number of participants was relatively small.

Table 1 lists the LDH Japan-affiliated artists and songs selected for both experiments. Shaded cells indicate experienced artists, while unshaded cells represent younger artists. The latest and representative songs shown in Table 1 were selected on October 30, 2024. For the representative songs, those of the experienced groups were independently selected by the author, based on each artist's popular tracks, while those of the younger groups were chosen as the songs with the highest number of views on each artist's official YouTube channel[2]. The artists are labeled A through F, the songs for each are labeled 1 through 3, and in Experiment 2, audio-only and MV conditions are distinguished by designating them as a or b, respectively. For example, the latest song (2) by GENERATIONS (C) is labeled as C2, and the MV (b) for "Lollipop" (3) by LIL LEAGUE (D) is labeled as D3b. Using these songs allows differences in the musical styles of artists from the same agency to be represented according to their debut periods and the selected songs. However, this approach does not control for differences arising from the time of composition or the year of release.

Table 1 Artists music in LDH Japan

Artist / Debut year	Experiment 1		Experiment 2
	Debut song (1)	Latest song (2)	Representative song (3)
EXILE (A) / 2001	Your eyes only	Reason	Rising sun (a, b)
J Soul Brothers III (B) / 2010	Best Friend's Girl	Baby don't cry	R. Y. U. S. E. I (a, b)
GENERATIONS (C) / 2012	BRAVE IT OUT	Endless·Journey	AGEHA (a, b)
LIL LEAGUE (D) / 2022	Hunter	Youth Spark	Lollipop (a, b)
WOLF HOWL HARMONY (E) / 2023	Sweet Rain	ROLLIN' STONES	Frozen Butterfly (a, b)
THE JET BOY BANGERZ (F) / 2023	Jettin'	UNBRAKABLE	PHOTOGENIC (a, b)

Table 2 presents the adjective pairs used in the SD method for both experiments. A total of 10 adjective pairs were chosen to investigate impressions related to the songs and MVs. Responses were rated on a 7-point scale, and the survey was conducted anonymously using Google Forms.

This study was conducted in accordance with the ethical guidelines for research involving human subjects at the Fukuoka Institute of Technology and received official approval (Approval

[2] https://www.youtube.com/@lilleague, https://www.youtube.com/@wolfhowlharmony,

https://www.youtube.com/@THE_JET_BOY_BANGERZ

No. hm18–24). Before participation, all participants were informed of the experimental procedures to be performed and precautions taken, and they provided voluntary informed consent.

Table 2 Adjective pairs for SD method

1 – 7	1 – 7	1 – 7
Powerful – Not powerful	Flashy – Modest	Fun – Sad
Passionate – Calm	Energetic – Quiet	Bright – Dark
Strong – Weak	Cool – Cute	Calm – Intense
		Like – Dislike

2.2 Experiment 1

The participants in the first experiment were divided into three age groups: 10 teenagers, 10 individuals in their twenties, 4 in their thirties, and 6 in their forties, totaling 30 participants (17 males and 13 females). Each participant was asked to evaluate their impressions of both the debut and latest songs of the artists listed in Table 1, using the 7-stage adjective scales shown in Table 2. We analyzed these results to determine how differences in artist generation and listener age influenced the impression ratings of the music.

2.3 Experiment 2

The participants were 22 individuals in their twenties (11 males and 11 females). Each participant was asked to evaluate their impressions of both the audio recordings and the MVs of the representative songs by the artists listed in Table 1, using the 7-point adjective scales shown in Table 2. We analyzed these results to demonstrate how the addition of visual information influenced the impression ratings compared to audio-only presentations.

3 EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Results of Experiment 1

Table 3 presents the correlation coefficients calculated from the mean ratings of each adjective pair. In Table 3, correlation coefficients greater than or equal to 0.8 are highlighted in red, whereas those less than or equal to −0.8 are highlighted in blue. The results indicate that strongly correlated adjective pairs were observed for all listener age groups. This suggests that some evaluation adjectives may be redundant, and therefore further refinement of the evaluation adjectives should be considered in future studies. Table 4 presents the communalities obtained from the factor analysis. In most cases, the communalities exceeded 0.7, indicating that the adjective pairs were well explained by the extracted common factors.

Table 5 presents the factor loadings that were extracted from the analysis. For all listener age groups, the number of factors was set to three based on the changes in eigenvalues that were observed using the scree plot method. The adjective pair “like–dislike” was excluded from the factor analysis, as it reflected evaluative preference. According to Table 5, the cumulative contribution rates for the three factors were 0.88, 0.94, and 0.91 for each age group. The factor loadings highlighted in red indicate the highest loading among the three factors for each adjective

pair. From these results, the three factors in the experiment were defined as shown in the second row of Table 5.

Table 3 Results of correlation matrix (Exp.1)

*Teens Twenties Thirty/Forties	Powerful–Not powerful	Passionate–Calm	Strong–Weak	Flashy–Modest	Energetic–Quiet	Cool–Cute	Fun–Sad	Bright–Dark	Calm–Intense
Powerful–Not powerful	1.00								
Passionate–Calm	0.82	1.00							
	0.80								
	0.76								
Strong–Weak	0.73	0.66	1.00						
	0.96	0.87							
	0.91	0.82							
Flashy–Modest	0.64	0.61	0.68	1.00					
	0.63	0.75	0.70						
	0.80	0.70	0.90						
Energetic–Quiet	0.77	0.60	0.71	0.92	1.00				
	0.79	0.78	0.85	0.87					
	0.82	0.69	0.87	0.92					
Cool–Cute	0.02	0.05	-0.06	-0.50	-0.50	1.00			
	0.18	-0.07	0.08	-0.37	-0.41				
	-0.47	-0.32	-0.49	-0.62	-0.67				
Fun–Sad	0.79	0.63	0.73	0.92	0.97	-0.42	1.00		
	0.76	0.78	0.83	0.79	0.98	-0.43			
	0.76	0.64	0.86	0.98	0.94	-0.70			
Bright–Dark	0.66	0.63	0.60	0.93	0.88	-0.37	0.93	1.00	
	0.67	0.79	0.79	0.79	0.94	-0.51	0.97		
	0.78	0.65	0.87	0.94	0.96	-0.74	0.96		
Calm–Intense	-0.89	-0.79	-0.73	-0.86	-0.93	0.21	-0.93	-0.87	1.00
	-0.93	-0.83	-0.94	-0.76	-0.94	0.11	-0.93	-0.84	
	-0.87	-0.52	-0.84	-0.74	-0.69	0.33	-0.70	-0.74	

Table 4 Results of communalities (Exp.1)

	Powerful–Not powerful	Passionate–Calm	Strong–Weak	Flashy–Modest	Energetic–Quiet	Cool–Cute	Fun–Sad	Bright–Dark	Calm–Intense
Teens	0.92	0.76	0.61	0.92	1.00	0.81	0.97	1.00	0.98
Twenties	0.99	0.75	0.95	1.00	0.99	0.85	1.00	0.95	0.98
Thirties/Forties	0.91	0.95	0.96	0.97	0.91	0.55	1.00	0.93	1.00

The results presented in Table 5 indicate that the responses from listeners in their teens and twenties showed factors that had largely similar tendencies. The contribution rates for each factor and the adjective pairs with high loadings were also comparable across these age groups. In contrast, the responses from listeners in their thirties and forties revealed distinct factors—such as *happiness*, *stability*, and *emotional depth*—that were not observed in the responses from the younger age groups. This may be because most participants in their teens and twenties were university students, whereas those in their thirties and forties were working adults. Since the songs that are commonly listened to during adolescence—a period when people are generally considered to listen to music most frequently—are likely to differ between current university students and working adults, their tendencies in Kansei evaluation of music may also differ.

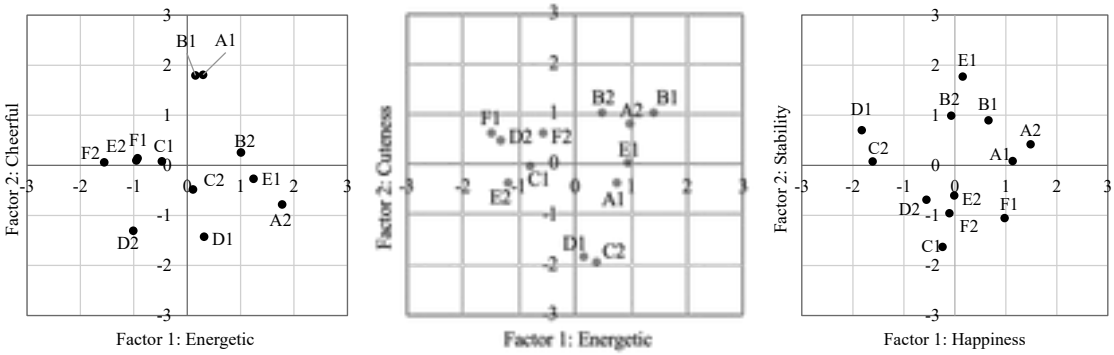
Table 5 Extracted factor loadings

Age group	Teens			Twenties			Thirties/Forties		
Extracted factors	F1: Energetic	F2: Cheerful	F3: Cuteness	F1: Energetic	F2 Cuteness	F3 Cheerful	F1: Happiness	F2: Stability	F3: Emotional depth
Contribution rate	0.46	0.26	0.17	0.61	0.20	0.14	0.40	0.29	0.23
Powerful–Not powerful	0.92	0.27		0.96	–0.17	0.23	0.36	0.70	0.53
Passionate–Calm	0.74	0.40	–0.22	0.77		0.39	0.28	0.27	0.89
Strong–Weak	0.72	0.26	0.13	0.93		0.27	0.50	0.61	0.58
Flashy–Modest	0.53	0.67	0.44	0.54	0.34	0.77	0.78	0.45	0.40
Energetic–Quiet	0.72	0.48	0.51	0.82	0.43	0.37	0.74	0.42	0.43
Cool–Cute		–0.19	–0.88		–0.91	–0.12	–0.73	–0.10	–0.11
Fun–Sad	0.68	0.60	0.39	0.84	0.50	0.23	0.85	0.40	0.33
Bright–Dark	0.49	0.83	0.27	0.74	0.57	0.27	0.77	0.47	0.34
Calm–Intense	–0.84	–0.49	–0.19	–0.94	–0.15	–0.27	–0.30	–0.93	–0.21

Figure 1 presents the mapping of factor scores for each song along with the first and second factors. Across all listener age groups, songs by experienced artists (A1, A2, B1, and B2) were positioned in proximity. In contrast, songs by younger artists (D1, D2, E1, E2, F1, and F2) tended to be located on the opposite side of the plot relative to the experienced songs. Table 6 gives the

mean response values for each song on the "like–dislike" adjective pair. In Table 6, the three lowest mean values are highlighted in red. These results indicate the latest songs tended to be rated as preferable than the debut songs.

Table 7 presents the clustering results based on the mean responses from the SD method for each song. Among listeners in their teens and twenties, the clustering patterns were largely consistent; however, participants in their thirties and forties classified D2 song into a different cluster. Cluster 1 primarily consisted of songs by experienced artists, and Clusters 2 and 3 were predominantly composed of songs by younger artists. The distinct clustering provided for song D2 suggests that artist D could evoke a different impression than artists E and F, despite also being categorized as a younger artist.



(a) Teens (b) Twenties (c) Thirties/Forties

Figure 1 Results of factor scores (Exp.1)

Table 6 Average values for "like – dislike"

	Teens	Twenties	Thirties/ Forties		Teens	Twenties	Thirties/ Forties
A1	3.00	2.11	3.33	A2	3.50	2.25	3.25
B1	3.63	2.00	3.13	B2	3.88	2.88	3.50
C1	3.38	2.13	3.00	C2	2.38	2.38	2.38
D1	1.88	2.25	2.88	D2	3.00	1.88	2.38
E1	3.50	2.25	2.63	E2	3.50	1.63	2.75
F1	3.25	2.25	2.75	F2	2.50	2.25	3.00

Table 7 Results of clustering (clusters: 3)

	Teens	Twenties	Thirties/Forties
Cluster 1	A1, B1, E1, A2, B2	A1, B1, E1, A2, B2	A1, B1, E1, A2, B2
Cluster 2	C1, F1, D2, E2, F2	C1, F1, D2, E2, F2	C1, F1, E2, F2
Cluster 3	D1, C2	D1, C2	D1, C2, D2

3.2 Results of Experiment 2

Table 8 presents the correlation matrix for the adjective pairs. Similar to Table 3, correlation coefficients greater than or equal to 0.8 (less than or equal to −0.8) are highlighted in red (blue). The results indicate a strong tendency for similar Kansei evaluations to be obtained across all songs and MVs. This suggests that the evaluations were similar because the evaluation objects consisted of only two types of stimuli, audio-only and visual (MV) information for each song and that no clear differences in evaluation tendencies attributable to the evaluation objects were observed among the participants. Table 9 presents the communalities obtained from the factor analysis. Similar to the Experiment 1, the communalities exceeded 0.7 in most cases, indicating that the adjective pairs were well explained by the extracted common factors.

Figure 2 presents the mapping of factor scores for each song on the first and second factors. The number of factors in the analysis was set to 2, based on the changes in eigenvalues that were observed using the scree plot method. The adjective pair “like–dislike” was excluded from the analysis because it represented evaluative preference. The contribution rates of the two factors were 0.75 and 0.19, respectively, and they had a cumulative contribution rate of 0.94. Because the factor loadings for each adjective pair closely matched the loadings for the responses of listeners in their twenties in Table 5, the factors were labeled identically: Factor 1 was labeled “Energetic” and Factor 2 “Cuteness.” From the data shown in Figure 2, it was clear that the impressions of the audio-only versions and the MVs of the same songs were largely similar.

Table 8 Results of correlation matrix (Exp.2)

	Powerful – Not powerful	Passionate – Calm	Strong – Weak	Flashy – Modest	Energetic – Quiet	Cool – Cute	Fun – Sad	Bright – Dark	Calm – Intense
Powerful – Not powerful	1.00								
Passionate – Calm	0.95	1.00							
Strong – Weak	0.95	0.99	1.00						
Flashy – Modest	0.94	0.90	0.91	1.00					
Energetic – Quiet	0.95	0.96	0.97	0.92	1.00				
Cool – Cute	0.43	0.58	0.61	0.41	0.55	1.00			
Fun – Sad	0.95	0.92	0.92	0.94	0.92	0.30	1.00		
Bright – Dark	0.93	0.96	0.96	0.93	0.96	0.52	0.96	1.00	
Calm – Intense	-0.91	-0.86	-0.89	-0.87	-0.94	-0.40	-0.86	-0.85	1.00

Table 9 Results of communalities (Exp.2)

Powerful– Not powerful	Passionate– Calm	Strong– Weak	Flashy– Modest	Energetic– Quiet	Cool– Cute	Fun– Sad	Bright– Dark	Calm– Intense
0.95	0.98	1.00	0.90	0.95	0.87	1.00	0.96	0.81

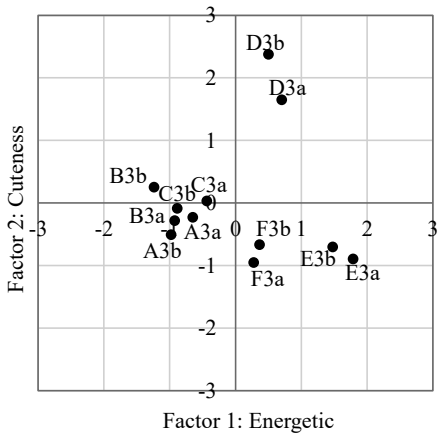


Figure 2 Results of factor scores (Exp.2)

3.3 Discussion

The results of Experiment 1 indicated that the factors extracted from the responses of listeners in their teens and twenties were noticeably different from those of listeners in their thirties and forties. The impressions of the songs may have varied across listener age groups. Factor analysis revealed that songs performed by experienced artists tended to be positioned more closely together, which indicates that listeners across age groups may have had similar impressions of these songs. Furthermore, the clustering results based on the mean responses from the SD method showed consistent classification patterns across all listener age groups. Therefore, the impressions of songs by experienced and younger artists were distinct.

The results of Experiment 2 confirmed that the audio-only and MV versions of the same songs elicited similar impression patterns. While it was hypothesized that the addition of visual information through MVs could alter participants' impressions, this trend was not observed in the experiment.

According to these findings, the results of Experiment 1 support verification item 1 as described in Section 1: there were clear differences in the factors that were extracted from the SD method responses between listeners in their teens and twenties and those in their thirties and forties. Experiment 1 also indicated that songs by experienced and younger artists produced distinct impression patterns, reflecting generational differences in the artists themselves. Finally, verification item 3 was addressed using the results of Experiment 2, which showed that the addition of visual information through MVs did not alter the participants' impressions of the songs. However, it cannot be excluded that the experimental results were affected not only by the limited participant sample but also by factors other than the artists' debut years, such as the musical characteristics of the songs.

3.4 Limitation & Future Research

It cannot be denied that the sample size in this study was relatively small; therefore, the results of the factor analysis should be regarded as exploratory. To examine the robustness of the findings, we conducted a leave-one-participant-out analysis by excluding each of the 10 participants in turn and performing factor analysis on the responses of the remaining nine

participants. The results showed that, for a particular exclusion case, a different factor structure was extracted compared with the other exclusion patterns. Therefore, future studies should include a larger number of participants to improve the robustness and generalizability of the findings. In addition, because the participants in their 30s and 40s were combined solely to balance the group sizes, it was not possible to examine age-related differences in Kansei evaluation among participants aged thirties and older.

Future research will focus on the relationship between musical features and impression ratings. This study focused solely on impression evaluations of songs and MVs, without taking into account the musical features themselves. If specific musical features influence user impressions, this will enable a more detailed analysis of interactions related to users' impression ratings. In addition, we will further investigate the differences in impression ratings between audio-only versions and MVs, including non-musical elements such as filming scenes related to the song's theme and the artist's clothing.

4 CONCLUSION

We conducted Kansei evaluations, taking into account the age of music listeners and generational differences among artists to analyze listeners' preference tendencies. The results showed that impressions of songs varied according to the listener's age group, and that songs by artists of different generations elicited distinct impressions. Moreover, it was suggested that the addition of visual information through MVs did not have an impact on impression ratings. In future, we will map musical attributes to Kansei evaluations to help clarify listener preferences.

REFERENCES

Hasegawa, P. R., Asakura, T. (2025). Reversal of Emotional Impressions through Key Change in Classical Music: An Assessment Using Major and Minor Transpositions. *International Journal of Affective Engineering*, 24(3), 267–277. https://doi.org/10.5057/ijae.IJAE-D-24-00052

Kakegawa, M., Komiyama, R., Masakura, Y., Kikuchi, M. (2012). Analysis of music appreciation by Kansei evaluation and brain activity. In *Proceedings of the 6th International Conference on Soft Computing and Intelligent Systems and the 13th International Symposium on Advanced Intelligence Systems*, (pp.183–186). https://doi.org/10.1109/SCIS-ISIS.2012.6505407

Kämpfe, J., Sedlmeier, P., & Renkewitz, F. (2010). The impact of background music on adult listeners: A meta-analysis. *Psychology of Music*, 39(4), 424–448. https://doi.org/10.1177/0305735610376261

Kurotaki, Y., Shiizuka, H. (2011). Measurement and Classification of the Affective Value of Drum Rhythm Patterns. In *Proceedings of 2011 Third International Conference on Intelligent Networking and Collaborative Systems*, (pp.872–875). https://doi.org/10.1109/INCoS.2011.116

Osgood, C. E. (1964). Semantic Differential Technique in the Comparative Study of Cultures. *American Anthropologist*, 66(3), 171–200. http://www.jstor.org/stable/669329