UNVEILING THE IMPACT OF MUSICAL FACTORS IN JUDGING A SONG ON FIRST LISTEN: INSIGHTS FROM A USER SURVEY

Kosetsu Tsukuda Tomoyasu Nakano Masahiro Hamasaki Masataka Goto
National Institute of Advanced Industrial Science and Technology (AIST), Japan
{k.tsukuda, t.nakano, masahiro.hamasaki, m.goto}@aist.go.jp

ABSTRACT

When a user listens to a song for the first time, what musical factors (e.g., melody, tempo, and lyrics) influence the user's decision to like or dislike the song? An answer to this question would enable researchers to more deeply understand how people interact with music. Thus, in this paper, we report the results of an online survey involving 302 participants to investigate the influence of 10 musical factors. We also evaluate how a user's personal characteristics (i.e., personality traits and musical sophistication) relate to the importance of each factor for the user. Moreover, we propose and evaluate three factor-based functions that would enable more effectively browsing songs on a music streaming service. The user survey results provide several reusable insights, including the following: (1) for most participants, the melody and singing voice are important factors in judging whether they like a song on first listen; (2) personal characteristics do influence the important factors (e.g., participants who have high openness and are sensitive to beat deviations emphasize melody); and (3) the proposed functions each have a certain level of demand because they enable users to easily find music that fits their tastes. We have released part of the survey results as publicly available data so that other researchers can reproduce the results and analyze the data from their own viewpoints.

1. INTRODUCTION

When a user listens to a song for the first time on a music streaming service and it matches her taste, she may listen to it until the end or add it to her favorites or a playlist. On the other hand, if the song does not match the user's preferences, she may stop playing it partway through [1,2]. By accumulating logs of such listening behaviors, music streaming services can estimate users' music preferences and implement functions such as recommendations [3,4].

However, when a user first listens to a song and decides whether or not she likes it, which musical factors influence the decision? For example, one user may like a song because of its lyrics, another may like it because of its melody, and third may like it because of the sound of a musical instrument. Several prior studies investigated people's preferred musical factors [5–7]. However, those studies targeted songs that the study participants already liked and investigated the reasons for liking those songs in terms of factors that were specific to the songs. Accordingly, when a participant answered that she liked a certain song because of its lyrics, it was unclear that she would always judge whether she liked or disliked a song because of its lyrics. Thus, despite those studies, there is a lack of research on the musical factors that influence people's judgment on whether they like a song on first listen. This lack of research motivates our first research question:

RQ1 When people listen to a song for the first time and judge whether they like it, which musical factors affect this judgment, and to what extent?

To more deeply understand how people interact with music, the effects of users' personality traits and musical sophistication on their music preferences and listening behaviors have also been studied [5, 8–21]. For example, it has been reported that people with high openness tend to show a preference for folk music [16] and that musical sophistication positively influences recommendation acceptance [20]. Following such studies, we address the second research question:

RQ2 How do people's personality traits and musical sophistication affect the importance of each musical factor in judging whether they like a song?

If a certain musical factor influences judgments about song preferences, it would be useful to propose practical examples of its engineering use. In fact, proposed improvements to the functions of music streaming services from user study results have provided useful insights to the music information retrieval (MIR) community [22–35]. Hence, we investigate a third research question:

RQ3 What are the implications of musical factors for the functions of music streaming services?

To address these research questions, we targeted 10 musical factors and conducted a questionnaire-based online user survey involving 302 participants. Our main contributions can be summarized as follows.

- We reveal that the factors of *melody* and *singing voice* have large influences on music preference judgment, whereas the factor of *danceability* has a small influence.
- From a psychological perspective, we show that both personality traits and musical sophistication affect the importance of the various musical factors. Given these results, we discuss the possibility that the important factors for a particular user could be estimated from the user's listening behaviors on a music streaming service.

- From an engineering perspective, we propose three functions that would enable users to effectively browse songs by leveraging musical factors, and we show that each function has a certain level of demand.
- We have made the English translation of the survey questionnaire and the survey results publicly available on the web to support future studies [1] .

2. RELATED WORK

2.1 Musical Factors

Understanding why people listen to music has been of interest to researchers. One typical research direction focuses on the motivation to listen to music in daily life. The main reasons include emotional reasons such as relaxation [18, 36–39] and relief [40, 41]. People also listen to music to concentrate and to pass time [42].

Another research direction investigates the reasons for listening to specific preferred songs in terms of musical factors. Greasley et al. [6] conducted interviews about participants' music collections. Among the main reasons why the participants liked their collections were musical factors such as the lyrics and instruments. Sanfilippo et al. [7] asked participants to sample two songs from their music library on a listening device and answer questions such as "why do you enjoy listening to the track?" The participants often answered the questions by using a vocabulary of musical factors. Boyle et al. [5] investigated the influence of musical factors on young people's pop music preferences. Each participant listed his/her three favorite pop songs and rated the importance of various musical factors in liking those songs. The results revealed that melody, mood, and rhythm had large influences. Although these studies investigated the influences of musical factors, they focused on only songs that the participants already liked. Our study is different in that we focus on the musical factors that people emphasize when they listen to a song for the first time. Since there is a vast number of songs that people have not yet listened to, investigating such factors is beneficial to support finding songs that match their preferences.

2.2 Personal Characteristics

In the music domain, user's preferences, interests, and behaviors are influenced by personal characteristics. In particular, many studies have investigated the influences of personality traits measured by the Big Five Inventory [8–14, 16, 17, 43–47]. For example, personality has significant associations with genre preferences [11, 13, 14, 16, 43] and audio preferences [47]. It also influences the desired level of diversity in a recommended song list [46]. Ferwerda et al. [45] revealed that when a user browses for music, the preferred taxonomy (mood, activity, and genre) depends on the user's personality. Such personality-based results can be used for personalization. In fact, several studies have shown increased recommendation quality when personality is incorporated [48–51]. Musical sophistication is another typical personal characteristic that influences music preferences. For example, musically sophisticated users listen to more diverse songs on both the artist and genre levels [52], are more familiar with the songs in a recommended song list [53], and prefer a less personalized playlist [19]. These findings can also be used to improve music recommendations and user interfaces. Following those studies, we investigate the influences of personality traits and musical sophistication on the importance of musical factors, and we suggest how its results can be used to improve the recommendations.

2.3 Design and Function Proposals

For user studies on music listeners' needs, preferences, and behaviors, it is common to not only report the results but also propose designs and functions to improve music services by applying the results [22–35]. Such proposals have provided reusable insights for the MIR community. Examples of these proposals include song recommendations according to the user's attention level [27], support for remote co-listening with a friend [31], and support for users to add their interpretations of lyrics [33]. Inspired by those prior studies, we propose three functions that enable music streaming services to leverage musical factors. Whereas the above studies only proposed designs and functions, we also conducted a user study to evaluate users' willingness to use the proposed functions.

3. PARTICIPANTS

We recruited participants for our user study via an online research company in Japan. We limited the participants to those who were Japanese and listened to music an average of at least one day per week via any music streaming service. The participants answered our questionnaire through a web browser. We paid about 13.21 USD (1,750 JPY) to each participant. Although 354 participants answered the survey, to make the analysis results more reliable, we removed the answers from 52 participants who submitted improper responses to a free-response question. The remaining 302 participants were diverse in both gender and age range: 147 male (10s: 4; 20s: 31; 30s: 33; 40s: 44; 50s: 35) and 155 female (10s: 9; 20s: 39; 30s: 35; 40s: 34; 50s: 38). Hereafter, we report the results obtained from the 302 participants including section 6.

4. INFLUENCE OF MUSICAL FACTORS

4.1 Musical Factors

Referring to prior studies on people's favorite songs [5–7, 54], we targeted the following 10 musical factors that may influence a person's judgment of liking or disliking music on first listen: *melody*, *singing voice*, *rhythm*, *lyrics*, *mood*, *tempo*, *harmony*, *sentiment*, *instruments*, and *danceability*. Although these 10 factors are not completely independent each other (e.g., there would be relatively high correlation between *mood* and *sentiment*), we adopted them to analyze as many factors as possible. In this study, all of these factors were determined entirely from the music. That is, we did not consider social factors that depend on the context of the music or the listener (e.g., the artist's image, the popularity of music, and whether music was introduced by a

[1] They can be downloaded from `https://github.com/ktsukuda/musical_factor`.

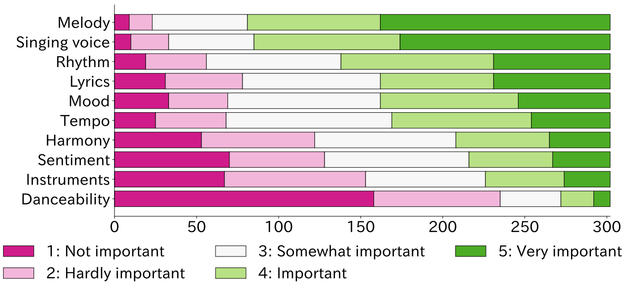

Figure 1. Importance distributions of musical factors (x-axis: number of participants).

friend). Rather, as this is an initial study on the influence of musical factors for judging a song on first listen, we leave the investigation of such social factors for future work.

4.2 Procedure

For each musical factor, we first showed the participants the factor's name, its meaning, and a question. In the case of *instruments*, for example, we showed the following description to represent its meaning: "*Instruments* means the type of instruments used in the piece and their sounds." Similarly, we showed the following question: "How important is the *instruments* in judging whether you like or dislike a song on first listen?" The possible answers were "not important," "hardly important," "somewhat important," "important," and "very important." When the answer for a factor was "not important" or "hardly important," the participant was asked to respond freely on why it was unimportant. On the other hand, when the answer was "somewhat important," "important," or "very important," the participant was asked to respond freely with at least one criterion for judging that he/she liked or disliked a song according to the factor. The 10 musical factors were displayed in a random order to each participant.

Note that in this survey, we asked the participants to answer the questions without actually listening to music to avoid answer bias caused by the music they listened to for the survey. Instead, they were asked to imagine daily situations where they listen to a song for the first time and rate the importance of each factor. This type of survey, which involves imagining a certain situation, is an established survey method in the MIR community [27, 31, 55–59].

4.3 Results

Figure 1 shows the importance distribution for each factor. We can see that the importance was high for *melody* and *singing voice*; in fact, paired Wilcoxon signed-rank tests with Bonferroni correction revealed that their medians (i.e., 4) were statistically higher than the medians of the remaining eight factors at $p < 0.01$. Among the remaining eight factors, more than half of the participants gave a rating of 3, 4, or 5 for *rhythm*, *lyrics*, *mood*, *tempo*, *harmony*, and *sentiment*. To more deeply understand the relationships between factors, we show the Spearman's rank correlations between them in Figure 2. There were high (> 0.4) correlations between *rhythm* and *tempo*, *mood* and *sentiment*, and *melody* and *singing voice*. Although *lyrics* had a relatively high average importance, it had low (< 0.3) correlations with all other factors. *Danceability*, which had

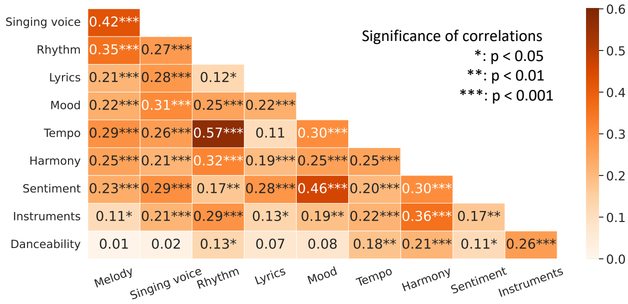

Figure 2. Spearman's rank correlations of importance between musical factors.

the lowest average importance, showed a similar tendency.

For each factor, to analyze the free responses on criteria for liking a song, we manually grouped the responses. Because we allowed the participants to give more than one criterion, each participant's response could be assigned to more than one group. Similarly, we grouped the responses on criteria for disliking a song and reasons for the unimportance of certain factors. Here, we omit the reasons for unimportance, because the most common response for all factors was "I am not interested in this factor." On the other hand, the criteria for liking or disliking a song were diverse, as seen in Table 1, which lists the top three criteria for each factor in terms of the group size. Many criteria involved opposite terms for liked and disliked songs: in the case of *tempo*, for example, participants who gave "fast" as a criterion for liking a song tended to give "slow" as a criterion for disliking a song. In addition, the second column indicates that, for all factors, more participants gave criteria for liking a song than for disliking a song, which means that it was more common to have criteria for liking a song than to have criteria for disliking a song. An interesting application of this finding would be to use criteria for liking a song in explainable recommendation. For example, when a song is recommended to a user who emphasizes *melody*, she may be more willing to listen to it if it appears with an explanation such as "this song is recommended to you because the melody is easy to remember."

The results in Figure 1 are somewhat similar to those reported by Boyle et al. [5] (e.g., *melody* and *rhythm* had high importance, while *danceability* had low importance). Nevertheless, we provide three contributions that are distinct from their results: (1) our results are more generalized, because we did not focus on a specific genre and age group, whereas they focused on young people's pop music preferences; (2) we analyzed the correlations between factors and the criteria for each factor; and (3) we will publish the survey results on the web to support later studies.

5. INFLUENCE OF PERSONAL FACTORS

5.1 Personality Traits

Procedure. We measured the participants' personality traits in terms of five aspects (i.e., *openness*, *conscientiousness*, *extraversion*, *agreeableness*, and *neuroticism*) by using the 29-item Big Five Inventory (BFI) on a 7-point scale (1: strongly disagree - 7: strongly agree) [60]. We used the BFI because of its popularity in past studies [8–14, 16, 17, 43–47] compared to other traits such as

Table 1. Top three criteria for judging "like" and "dislike," for each musical factor. Each number in parentheses indicates the number of participants who responded with the corresponding criterion.

Factor		1st	2nd	3rd
Melody	Like (265)	Easy to remember (35)	Easy to sing or hum (33)	Feels comfortable (28)
	Dislike (193)	Too loud (18)	Difficult to sing or hum (16)	Feels uncomfortable (15)
Singing voice	Like (261)	Specific type (beautiful, powerful, soft, etc.) (74)	Voice to my liking (54)	Feels comfortable (51)
	Dislike (203)	Feels uncomfortable (50)	Specific type (raspy, piercing, etc.) (47)	Voice not to my liking (28)
Rhythm	Like (237)	Groovy (53)	Feels comfortable (23)	Rhythm to my liking (19)
	Dislike (167)	Rhythm not to my liking (17)	Slow (16)	Not groovy (15)
Lyrics	Like (218)	Sympathetic (71)	Inspirational (41)	Positive (10)
	Dislike (164)	Unclear meaning (41)	Lack empathy (30)	Pedestrian (26)
Mood	Like (219)	Cheerful (51)	Fits my mood/situation (25)	Calm (21)
	Dislike (162)	Gloomy (32)	Too loud (29)	Feels uncomfortable (12)
Tempo	Like (220)	Fast (40)	Groovy (29)	Feels comfortable (24)
	Dislike (163)	Slow (48)	Fast (31)	Feels uncomfortable (15)
Harmony	Like (174)	Feels comfortable (43)	Beautiful (23)	Harmonious (22)
	Dislike (116)	Feels uncomfortable (25)	Monotonous (7)	Inharmonious (6)
Sentiment	Like (163)	Positive (33)	Inspirational (30)	Sympathetic (25)
	Dislike (114)	Negative (32)	Evokes no emotion (12)	Doesn't fit my mood/situation (7)
Instruments	Like (146)	Include specific instruments (24)	Fit the song (17)	Feel comfortable (15)
	Dislike (102)	Too loud (24)	Feel uncomfortable (11)	Don't fit the song (7)
Danceability	Like (66)	Body moves naturally to music (13)	Groovy (11)	Rhythmic (9)
	Dislike (46)	Not groovy (6)	Gloomy (5)	Rhythm is bad (4)

Table 2. Spearman's rank correlations between personality traits and musical factor importance (N=302). Significant correlations are shown in bold (*: p$<$ 0.05; **: p$<$ 0.01; ***: p$<$ 0.001).

Trait	Melody	Singing voice	Rhythm	Lyrics	Mood	Tempo	Harmony	Sentiment	Instruments	Danceability
Openness	**0.127***	**0.135***	**0.155****	**0.177****	0.107	0.109	**0.255*****	0.050	**0.157****	**0.151****
Conscientiousness	0.076	**0.128***	0.062	0.031	**0.127***	**0.128***	**0.125***	**0.119***	0.028	0.013
Extraversion	0.062	**0.130***	**0.172****	**0.175****	0.098	**0.114***	**0.254*****	0.107	**0.151****	**0.219*****
Agreeableness	0.025	**0.123***	0.048	0.088	**0.158****	0.029	0.021	0.060	0.049	0.065
Neuroticism	0.003	0.010	-0.081	0.036	-0.025	-0.004	**-0.142***	0.109	-0.072	**-0.120***

opinion leadership [15].

Results. Table 2 lists the Spearman's rank correlations between the personality traits and the importance of each musical factor. *Openness* had significant correlations with as many as seven factors. That is, participants with higher *openness* had more diverse criteria for judging whether a song fits their taste. This result is similar to a previous finding that people with high *openness* tended to listen to more diverse songs in terms of genres [16]. Similarly, *extraversion* also had significant correlations with many factors, particularly, *danceability*. This result echoes a report that people with high *extraversion* tended to listen to songs with high danceability on a music streaming service [51]. *Conscientiousness* was the only trait that had a significant correlation with *sentiment*. Both *agreeableness* and *neuroticism* had significant correlations with as few as two factors. These results are similar to a previous finding that those traits showed significant correlations with few genres [16].

Prior studies correlated personality traits with genre preferences and music audio preferences [16, 47]. For example, people who often listen to folk music were found to have high *openness* [16]. As seen in Table 2, people with high *openness* emphasize *lyrics*; accordingly, for a user who often listens to folk songs, it would be helpful to recommend songs according to the similarity of lyrics.

5.2 Musical Sophistication

Procedure. To measure the musical sophistication, we used the following nine questions on a 7-point scale.

1. InstExp: I engage in regular, daily practice of a musical instrument (1: never - 7: $\geq$ 10 years).
2. DanceExp: I engage in regular, daily dancing (1: never - 7. more than 10 years).
3. NoticeBeat: I can tell when people sing or play out of time with the beat (1: strongly disagree - 7: strongly agree).
4. NoticeTune: I can tell when people sing or play out of tune (1: strongly disagree - 7: strongly agree).
5. LsnMusic: I listen to music (1: $<$ 15 minutes per day - 7: $\geq$ 4 hours per day).
6. LsnNew: I listen to music that is new to me (1: $<$ 1 song per month - 7: $\geq$ 31 songs per month).
7. ViewLyrics: I view lyrics while listening to music (1: $<$ 1 song per month - 7: $\geq$ 31 songs per month).
8. Karaoke: I sing karaoke (1: $<$ 1 time per year - 7: $\geq$ 4 times per week).
9. AttEvt: I attend live music events as an audience member (1: $<$ 1 time per year - 7: $\geq$ 11 times per year).

Questions 1, 3, 4, 5, and 9 derive from the Goldsmiths Musical Sophistication Index (Gold-MSI) [61]. In addition, we asked four questions of our own (questions 2, 6, 7, and 8). For questions 5-9, we asked the participants to give the average frequencies of those behaviors.

Results. Table 3 lists the Spearman's rank correlations between musical sophistication and the importance of each musical factor. Overall, many of the results matched our intuition. For example, DanceExp had a significantly high correlation with *danceability*; participants who were sensitive to beat and tune deviations emphasized audio-based factors such as *melody*, *singing voice*, and *harmony*; and ViewLyrics had the highest correlation with *lyrics*. It is also convincing that participants who often sang karaoke emphasized *lyrics*; those who often attended live music

Table 3. Spearman's rank correlations between musical sophistication and the importance of each musical factor (N=302). Significant correlations are shown in bold (*: $p< 0.05$; **: $p< 0.01$; ***: $p< 0.001$).

Question	Melody	Singing voice	Rhythm	Lyrics	Mood	Tempo	Harmony	Sentiment	Instruments	Danceability
InstExp	0.100	0.061	-0.019	0.108	0.037	-0.099	**0.134***	0.093	0.091	0.101
DanceExp	-0.041	0.039	-0.047	**0.126***	0.030	-0.024	0.044	0.098	-0.005	**0.341*****
NoticeBeat	**0.228*****	**0.228*****	**0.126***	0.082	0.107	0.073	**0.302*****	**0.205*****	**0.147***	0.072
NoticeTune	**0.272*****	**0.231*****	0.099	0.088	**0.121***	0.039	**0.276*****	**0.167****	0.078	0.001
LsnMusic	0.041	0.054	0.111	**0.141***	**0.135***	0.101	0.078	0.108	0.051	0.090
LsnNew	0.003	0.107	**0.152****	**0.152****	0.112	**0.194*****	**0.115***	0.101	**0.126***	**0.169****
ViewLyrics	0.001	0.085	**0.118***	**0.243*****	**0.120***	**0.147***	**0.136***	**0.128***	0.101	0.110
Karaoke	0.085	0.087	0.005	**0.210*****	**0.154****	-0.015	0.057	**0.129***	-0.033	0.081
AttEvent	-0.038	0.037	-0.023	**0.200*****	0.004	0.016	0.039	-0.005	0.088	**0.179****

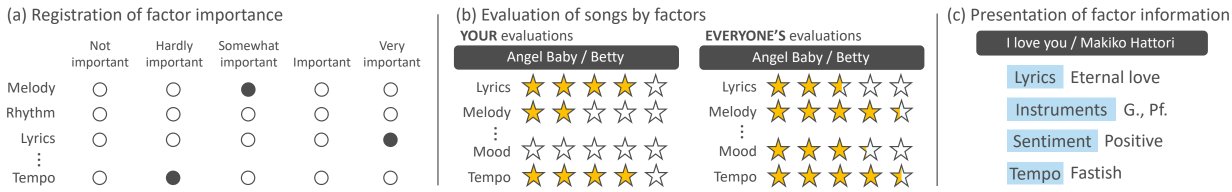

Figure 3. Overview of the three proposed functions. In the user study, these images were presented to the participants.

events emphasized both *lyrics* and *danceability*; and InstExp had a significant correlation with *harmony*. Table 3 also indicates certain high correlations that are not obvious (e.g., between LsnMusic/LsnNew and *lyrics* and between LsnNew and *danceability*).

Certain metrics, such as LsnMusic, LsnNew, and ViewLyrics, can be computed for each user on a music streaming service [59,62,63]. Thus, the results in Table 3 can also be used to increase the confidence in estimating the importance of each factor to a user without explicitly asking the importance. For example, if a user often listens to folk music (i.e., the user would have high *openness* as has been reported by Ferwerda et al. [16]) and new songs, we can estimate from the results in Tables 2 and 3 that *rhythm* is one of the user's important factors. Hence, the user would be more likely to accept recommendations by recommending songs according to the similarity of their rhythms.

6. FUNCTIONS BASED ON MUSICAL FACTORS

In section 4, we showed that certain musical factors influence a person's judgment of liking or disliking a song on first listen. Following those results, in this section, we propose three functions, illustrated in Figure 3, that could enrich and diversify the music listening experience on streaming services. Then, we investigate the usefulness of these functions from the results of a user study.

6.1 Functions

6.1.1 Function 1: Registration of Factor Importance

With this function, shown in Figure 3 (a), users register the importance of each of the 10 musical factors on a 5-point scale when judging whether they like or dislike music on first listen. It is not necessary to register the importance of all factors. For example, the importance of *rhythm* is not registered in Figure 3 (a). The registration process only needs to be done once, and the registered information can be changed later.

This function supports the users as follows. Suppose that a user is listening to her favorite song s. The user has registered *lyrics* as "very important" and *tempo* as "hardly important." Hence, among songs that are new to this user, we can recommend songs that have various tempos and similar lyrics to s. By listening to the recommended songs, the user can find new favorite songs.

6.1.2 Function 2: Evaluation of Songs by Factors

This proposed function allows users to rate their song preferences on a factor-by-factor basis, as shown in Figure 3 (b). The ratings are not mandatory: users only need to rate the songs that they want to rate. In addition, they do not need to rate songs in terms of all 10 factors. For example, in the figure, the user does not rate *mood*. For each song, by computing the average value of all users' rating results for each factor, we can display others' evaluations (averaged ratings) like those shown in Figure 3 (b).

This function supports the users as follows. Suppose that a user is interested in an artist named "Betty," and that *danceability* is an important factor for the user. Then, songs by "Betty" can be sorted and displayed in order of the averaged ratings for *danceability*. This enables efficient discovery of songs that match the user's preferences.

6.1.3 Function 3: Presentation of Factor Information

With this function, information on factors that a user wants to know for a song is displayed as shown in Figure 3 (c). The information on each of the 10 factors can be automatically estimated by using techniques from existing studies [64–70]. Thus, unlike the two previous functions, this one does not require the user to input any information.

This function supports the users as follows. When a user checks a list of newly released songs, usually only basic information such as the artist and title is displayed for each song. In contrast, our proposed function can display information on the musical factor for each song. For example, if the user prefers slow-tempo songs with piano, she can listen only to such songs by referring to the displayed information on *tempo* and *instruments*. This allows the user to efficiently find songs that match her preferences among a vast number of new songs.

Table 4. Top three free-response reasons for "reasonably willing" or "willing" to use each of the proposed functions. Each number in parentheses indicates the number of participants who gave that reason.

	Function 1: registration of factor importance	Function 2: evaluation of songs by factors	Function 3: presentation of factor information
1st	Easy to find music that fits my taste. (46)	Would like to refer to others' evaluations. (22)	Easy to find music that fits my mood/situation. (27)
2nd	Helpful for listening to new songs. (33)	Easy to understand others' evaluations. (14)	Easy to find music that fits my taste. (26)
3rd	Looks interesting to use. (11)	Easy to find music that fits my taste. (13)	Helpful for listening to new songs. (16)

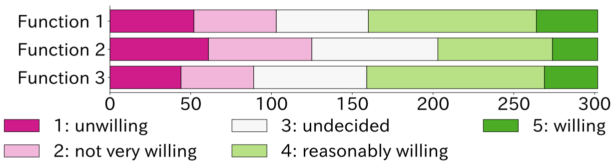

Figure 4. Distribution of the willingness to use each of the proposed functions (x-axis: number of participants).

6.2 Procedure

For each function, we showed the participants an overview of the function and examples of the user support that the function would enable as we described in section 6.1 [2]. The participants were asked to indicate their willingness to use the function, on a 5-point scale ("unwilling," "not very willing," "undecided," "reasonably willing," and "willing"), if it were implemented on the music streaming service that they used regularly. They were also asked to provide free responses on their willingness. The three functions were displayed in a random order to each participant.

6.3 Results

Figure 4 shows the answer distribution for each function. Functions 1 and 3 were more positively received than function 2. To analyze the results, we manually grouped negative responses (i.e., the free responses for "unwilling" and "not very willing"). As we had anticipated, a reason of "I do not need the function" was common for all three functions. Regarding function 2, although we explained that the ratings were not mandatory, a response of "It is tedious to rate songs" was also common. This is why the distribution for function 2 was more biased in the negative direction. Here, note that our goal was not to propose functions that all participants would be willing to use. Rather, we sought to confirm that the proposed functions would have a certain level of demand; accordingly, the results in Figure 4 indicate that we achieved our objective.

We also manually grouped the positive responses (i.e., the free responses for "willing" and "reasonably willing"). Table 4 lists the top three responses in terms of the group size for each function. We can see that, in general, the participants tended to appreciate functions that would make it easy to find music that fits their taste (all functions) and easy to listen to new songs (functions 1 and 3). The responses for function 2 also indicate that they were interested in referring to other users' evaluations of a song. We can also see that the participants felt it was valuable to be able to find music according to their mood or situation (function 3). These responses provide reusable insights for later studies: when researchers or streaming services propose a new function, such user demand could serve as a useful guideline for its design.

If function 3 were implemented on a music streaming service, it might be difficult to estimate the information for all factors because of the platform's resource limitations. In such a case, a possible solution would be to decrease the number of displayed factors according to the results shown in Figure 2. For example, *rhythm* information could be omitted, because *tempo* has a high correlation with *rhythm*, and users who emphasize *rhythm* could thus refer to *tempo* information instead. In contrast, *lyrics* should not be eliminated because it has low correlations with the other factors, and there would not be no alternative factor for users who emphasize *lyrics*.

7. CONCLUSION

In this paper, we conducted an online user survey involving 302 participants. The reusable insights obtained from our user survey can be summarized as follows.

- We showed that the *melody* and *singing voice* are important for most participants. Because there were trends in the criteria for each factor, as seen in Table 1, the criteria could be used to increase the explainability of song recommendations, as discussed in section 4.3.
- Personality and musical sophistication influence the importance of each musical factor. As discussed in sections 5.1 and 5.2, these results would be useful for estimating which factors are important to a user from the user's listening behaviors on a streaming service.
- The evaluation results for our proposed functions show that there is a certain demand for functions that enable users to browse songs according to musical factors. The reasons for each function's demand in Table 4 could provide guidelines for other researchers and services to propose novel factor-based functions.

Finally, we acknowledge a limitation of this paper in that all the participants in our user study were Japanese. Because peoples' music preferences and listening behaviors, as well as music itself, vary widely from country to country [26, 71–76], not all of the findings reported here can be generalized. Nevertheless, we believe that our study provides a worthwhile contribution to the MIR community as a first step toward understanding how musical factors influence whether people like a song on first listen. At the same time, the above limitation can guide future work such as investigating the differences in important musical factors among countries and cultures. The publicly available dataset of results from our user study will enable researchers not only to perform such comparisons but also to analyze and compare results from different viewpoints.

[2] We leave it as future work to actually implement these functions and conduct a long-term user study on them including how to visualize the information.

8. ACKNOWLEDGMENTS

This work was supported in part by JST CREST Grant Number JPMJCR20D4 and JSPS KAKENHI Grant Number JP21H04917, Japan.

9. REFERENCES

[1] H. Yakura, T. Nakano, and M. Goto, "FocusMusicRecommender: A system for recommending music to listen to while working," in *Proceedings of the 23rd International Conference on Intelligent User Interfaces*, ser. IUI 2018, 2018, pp. 7–17.

[2] B. Brost, R. Mehrotra, and T. Jehan, "The music streaming sessions dataset," in *Proceedings of the World Wide Web Conference*, ser. WWW 2019, 2019, pp. 2594–2600.

[3] P. Knees and M. Schedl, "A survey of music similarity and recommendation from music context data," *ACM Transactions on Multimedia Computing, Communications, and Applications*, vol. 10, no. 1, pp. 1–21, 2013.

[4] Y. Deldjoo, M. Schedl, P. Cremonesi, and G. Pasi, "Recommender systems leveraging multimedia content," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–38, 2020.

[5] J. D. Boyle, G. L. Hosterman, and D. S. Ramsey, "Factors influencing pop music preferences of young people," *Journal of Research in Music Education*, vol. 29, no. 1, pp. 47–55, 1981.

[6] A. Greasley, A. Lamont, and J. A. Sloboda, "Exploring musical preferences: An in-depth qualitative study of adults' liking for music in their personal collections," *Qualitative Research in Psychology*, vol. 10, no. 4, pp. 402–427, 2013.

[7] K. R. M. Sanfilippo, N. Spiro, M. Molina-Solana, and A. Lamont, "Do the shuffle: Exploring reasons for music listening through shuffled play," *PLOS ONE*, vol. 15, no. 2, pp. 1–21, 2020.

[8] T. Chamorro-Premuzic and A. Furnham, "Personality and music: Can traits explain how people use music in everyday life?" *British Journal of Psychology*, vol. 98, no. 2, pp. 175–185, 2007.

[9] M. J. Delsing, T. F. Ter Bogt, R. C. Engels, and W. H. Meeus, "Adolescents' music preferences and personality characteristics," *European Journal of Personality*, vol. 22, no. 2, pp. 109–130, 2008.

[10] R. L. Zweigenhaft, "A do re mi encore: A closer look at the personality correlates of music preferences," *Journal of individual differences*, vol. 29, no. 1, pp. 45–55, 2008.

[11] R. A. Brown, "Music preferences and personality among japanese university students," *International Journal of Psychology*, vol. 47, no. 4, pp. 259–268, 2012.

[12] T. Chamorro-Premuzic, V. Swami, and B. Cermakova, "Individual differences in music consumption are predicted by uses of music and age rather than emotional intelligence, neuroticism, extraversion or openness," *Psychology of Music*, vol. 40, no. 3, pp. 285–300, 2012.

[13] A. Langmeyer, A. Guglhör-Rudan, and C. Tarnai, "What do music preferences reveal about personality?" *Journal of Individual Differences*, vol. 33, no. 2, pp. 119–130, 2012.

[14] A. Laplante, "Improving music recommender systems: What can we learn from research on music tastes?" in *Proceedings of the 15th International Society for Music Information Retrieval Conference*, ser. ISMIR 2014, 2014, pp. 451–456.

[15] A. E. Krause and A. C. North, "Music listening in everyday life: Devices, selection methods, and digital technology," *Psychology of Music*, vol. 44, no. 1, pp. 129–147, 2016.

[16] B. Ferwerda, M. Tkalcic, and M. Schedl, "Personality traits and music genres: What do people prefer to listen to?" in *Proceedings of the 25th ACM Conference on User Modeling, Adaptation and Personalization*, ser. UMAP 2017, 2017, pp. 285–288.

[17] T. Schäfer and C. Mehlhorn, "Can personality traits predict musical style preferences? A meta-analysis," *Personality and Individual Differences*, vol. 116, pp. 265–273, 2017.

[18] W. M. Randall and N. S. Rickard, "Reasons for personal music listening: A mobile experience sampling study of emotional outcomes," *Psychology of Music*, vol. 45, no. 4, pp. 479–495, 2017.

[19] Y. Liang and M. C. Willemsen, "Personalized recommendations for music genre exploration," in *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization*, ser. UMAP 2019, 2019, pp. 276–284.

[20] Y. Jin, N. Tintarev, and K. Verbert, "Effects of personal characteristics on music recommender systems with different levels of controllability," in *Proceedings of the 12th ACM Conference on Recommender Systems*, ser. RecSys 2018, 2018, pp. 13–21.

[21] Y. Liang and M. C. Willemsen, "The role of preference consistency, defaults and musical expertise in users' exploration behavior in a genre exploration recommender," in *Proceedings of the 15th ACM Conference on Recommender Systems*, ser. RecSys 2021, 2021, pp. 230–240.

[22] J. S. Downie and S. J. Cunningham, "Toward a theory of music information retrieval queries: System design implications," in *Proceedings of the 3rd International Conference on Music Information Retrieval*, ser. ISMIR 2002, 2002, pp. 299–300.

[23] S. J. Cunningham, N. Reeves, and M. Britland, "An ethnographic study of music information seeking: Implications for the design of a music digital library," in *Proceedings of the 3rd ACM/IEEE-CS Joint Conference on Digital Libraries*, ser. JCDL 2003, 2003, pp. 5–16.

[24] S. Jones, S. J. Cunningham, and M. Jones, "Organizing digital music for use: An examination of personal music collections," in *Proceedings of the 5th International Conference on Music Information Retrieval*, ser. ISMIR 2004, 2004, pp. 397–402.

[25] C. Inskip, R. Butterworth, and A. MacFarlane, "A study of the information needs of the users of a folk music library and the implications for the design of a digital library system," *Information Processing & Management*, vol. 44, no. 2, pp. 647–662, 2008.

[26] X. Hu, J. H. Lee, and L. K. Y. Wong, "Music information behaviors and system preferences of university students in Hong Kong," in *Proceedings of the 15th International Society for Music Information Retrieval Conference*, ser. ISMIR 2014, 2014, pp. 579–584.

[27] J. H. Lee and R. Price, "Understanding users of commercial music services through personas: Design implications," in *Proceedings of the 16th International Society for Music Information Retrieval Conference*, ser. ISMIR 2015, 2015, pp. 476–482.

[28] J. H. Lee, H. Cho, and Y.-S. Kim, "Users' music information needs and behaviors: Design implications for music information retrieval systems," *Journal of the Association for Information Science and Technology*, vol. 67, no. 6, pp. 1301–1330, 2016.

[29] J. H. Lee, Y. Kim, and C. Hubbles, "A look at the cloud from both sides now: An analysis of cloud music service usage," in *Proceedings of the 17th International Society for Music Information Retrieval Conference*, ser. ISMIR 2016, 2016, pp. 299–305.

[30] L. Spinelli, J. Lau, L. Pritchard, and J. H. Lee, "Influences on the social practices surrounding commercial music services: A model for rich interactions," in *Proceedings of the 19th International Society for Music Information Retrieval Conference*, ser. ISMIR 2018, 2018, pp. 671–677.

[31] J. H. Lee, L. Pritchard, and C. Hubbles, "Can we listen to it together?: Factors influencing reception of music recommendations and post-recommendation behavior," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, ser. ISMIR 2019, 2019, pp. 663–669.

[32] J. H. Lee and A. T. Nguyen, "How music fans shape commercial music services: A case study of BTS and ARMY," in *Proceedings of the 21st International Society for Music Information Retrieval Conference*, ser. ISMIR 2020, 2020, pp. 837–845.

[33] J. H. Lee, A. Bhattacharya, R. Antony, N. Santero, and A. Le, ""Finding home": Understanding how music supports listeners' mental health through a case study of BTS," in *Proceedings of the 22nd International Society for Music Information Retrieval Conference*, ser. ISMIR 2021, 2021, pp. 358–365.

[34] X. Hu, J. Chen, and Y. Wang, "University students' use of music for learning and well-being: A qualitative study and design implications," *Information Processing & Management*, vol. 58, no. 1, pp. 1–14, 2021.

[35] S. Y. Park and B. Kaneshiro, "Social music curation that works: Insights from successful collaborative playlists," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW1, pp. 1–27, 2021.

[36] J. A. Sloboda, S. A. O'Neill, and A. Ivaldi, "Functions of music in everyday life: An exploratory study using the experience sampling method," *Musicae Scientiae*, vol. 5, no. 1, pp. 9–32, 2001.

[37] A. Lamont and R. Webb, "Short- and long-term musical preferences: What makes a favourite piece of music?" *Psychology of Music*, vol. 38, no. 2, pp. 222–241, 2010.

[38] A. B. Haake, "Individual music listening in workplace settings: An exploratory survey of offices in the UK," *Musicae Scientiae*, vol. 15, no. 1, pp. 107–129, 2011.

[39] T. Schäfer, "The goals and effects of music listening and their relationship to the strength of music preference," *PloS ONE*, vol. 11, no. 3, pp. 1–15, 2016.

[40] A. J. Lonsdale and A. C. North, "Why do we listen to music? A uses and gratifications analysis," *British Journal of Psychology*, vol. 102, no. 1, pp. 108–134, 2011.

[41] S. Y. Park, E. Redmond, J. Berger, and B. Kaneshiro, "Hitting pause: How user perceptions of collaborative playlists evolved in the united states during the covid-19 pandemic," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI 2022, 2022, pp. 1–16.

[42] A. C. North, D. J. Hargreaves, and J. J. Hargreaves, "Uses of music in everyday life," *Music Perception: An Interdisciplinary Journal*, vol. 22, no. 1, pp. 41–77, 2004.

[43] P. J. Rentfrow and S. D. Gosling, "The do re mi's of everyday life: The structure and personality correlates of music preferences." *Journal of Personality and Social Psychology*, vol. 84, no. 6, pp. 1236–1256, 2003.

[44] M. Tkalčič, B. Ferwerda, D. Hauger, and M. Schedl, "Personality correlates for digital concert program notes," in *Proceedings of the 23rd ACM Conference on User Modeling, Adaptation and Personalization*, ser. UMAP 2015, 2015, pp. 364–369.

[45] B. Ferwerda, E. Yang, M. Schedl, and M. Tkalcic, "Personality traits predict music taxonomy preferences," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA 2015, 2015, pp. 2241–2246.

[46] B. Ferwerda, M. Graus, A. Vall, M. Tkalcic, and M. Schedl, "The influence of users' personality traits on satisfaction and attractiveness of diversified recommendation lists," in *Proceedings of the 4th Workshop on Emotions and Personality in Personalized Systems*, ser. EMPIRE 2016, 2016, pp. 43–47.

[47] A. B. Melchiorre and M. Schedl, "Personality correlates of music audio preferences for modelling music listeners," in *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, ser. UMAP 2020, 2020, pp. 313–317.

[48] R. Hu and P. Pu, "Enhancing collaborative filtering systems with personality information," in *Proceedings of the 5th ACM Conference on Recommender Systems*, ser. RecSys 2011, 2011, pp. 197–204.

[49] I. Fernández-Tobías, M. Braunhofer, M. Elahi, F. Ricci, and I. Cantador, "Alleviating the new user problem in collaborative filtering by exploiting personality information," *User Modeling and User-Adapted Interaction*, vol. 26, no. 2, pp. 221–255, 2016.

[50] F. Lu and N. Tintarev, "A diversity adjusting strategy with personality for music recommendation." in *Proceedings of the 5th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems*, ser. IntRS 2018, 2018, pp. 7–14.

[51] A. B. Melchiorre, E. Zangerle, and M. Schedl, "Personality bias of music recommendation algorithms," in *Proceedings of the 14th ACM Conference on Recommender Systems*, ser. RecSys 2020, 2020, pp. 533–538.

[52] B. Ferwerda and M. Tkalčič, "Exploring online music listening behaviors of musically sophisticated users," in *Proceedings of the 27th Conference on User Modeling, Adaptation and Personalization*, ser. UMAP 2019, 2019, pp. 33–37.

[53] B. Ferwerda, M. P. Graus, A. Vall, M. Tkalcic, and M. Schedl, "How item discovery enabled by diversity leads to increased recommendation list attractiveness," in *Proceedings of the 32nd ACM SIGAPP Symposium on Applied Computing*, ser. SAC 2017, 2017, pp. 1693–1696.

[54] A. LeBlanc, "Outline of a proposed model of sources of variation in musical taste," *Bulletin of the Council for Research in Music Education*, no. 61, pp. 29–34, 1980.

[55] J. H. Lee and J. S. Downie, "Survey of music information needs, uses, and seeking behaviours: Preliminary findings," in *Proceedings of the 5th International Conference on Music Information Retrieval*, ser. ISMIR 2004, 2004, pp. 441–446.

[56] A. Laplante, "Users' relevance criteria in music retrieval in everyday life: An exploratory study," in *Proceedings of the 11th International Society for Music Information Retrieval Conference*, ser. ISMIR 2010, 2010, pp. 601–606.

[57] J. H. Lee and N. M. Waterman, "Understanding user requirements for music information services," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 253–258.

[58] M. Kamalzadeh, D. Baur, and T. Möller, "A survey on music listening and management behaviours," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 373–378.

[59] K. Tsukuda, M. Hamasaki, and M. Goto, "Toward an understanding of lyrics-viewing behavior while listening to music on a smartphone," in *Proceedings of the 22nd International Society for Music Information Retrieval Conference*, ser. ISMIR 2021, 2021, pp. 705–713.

[60] T. Namikawa, I. Tani, T. Wakita, R. Kumagai, A. Nakane, and H. Noguchi, "Development of a short form of the Japanese Big-Five Scale, and a test of its reliability and validity," *The Japanese Journal of Psychology*, vol. 83, no. 2, pp. 91–99, 2012.

[61] D. Müllensiefen, B. Gingras, J. Musil, and L. Stewart, "The musicality of non-musicians: An index for assessing musical sophistication in the general population," *PLOS ONE*, vol. 9, no. 2, pp. 1–23, 2014.

[62] G. Vigliensoni and I. Fujinaga, "The music listening histories dataset," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ser. ISMIR 2017, 2017, pp. 96–102.

[63] M. Schedl, S. Brandl, O. Lesota, E. Parada-Cabaleiro, D. Penz, and N. Rekabsaz, "LFM-2b: A dataset of enriched music listening events for recommender systems research and fairness analysis," in *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*, ser. CHIIR 2022, 2022, pp. 337–341.

[64] K. Tsukuda, K. Ishida, and M. Goto, "Lyric Jumper: A lyrics-based music exploratory web service by modeling lyrics generative process," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ser. ISMIR 2017, 2017, pp. 544–551.

[65] R. Delbouys, R. Hennequin, F. Piccoli, J. Royo-Letelier, and M. Moussallam, "Music mood detection based on audio and lyrics with deep neural net," in *Proceedings of the 19th International Society for Music*

Information Retrieval Conference, ser. ISMIR 2018, 2018, pp. 370–375.

[66] F. Karsdorp, P. van Kranenburg, and E. Manjavacas, "Learning similarity metrics for melody retrieval," in *Proceedings of the 20th International Society for Music Information Retrieval Conference*, ser. ISMIR 2019, 2019, pp. 478–485.

[67] M. Hamasaki, K. Ishida, T. Nakano, and M. Goto, "Songrium RelayPlay: A web-based listening interface for continuously playing user-generated music videos of the same song with different singers," in *Proceedings of the International Computer Music Conference 2020*, ser. ICMC 2020, 2020, pp. 426–429.

[68] A. A. Correya, D. Bogdanov, L. Joglar-Ongay, and X. Serra, "Essentia.js: A javascript library for music and audio analysis on the web," in *Proceedings of the 21st International Society for Music Information Retrieval Conference*, ser. ISMIR 2020, 2020, pp. 605–612.

[69] G. Micchi, K. Kosta, G. Medeot, and P. Chanquion, "A deep learning method for enforcing coherence in automatic chord recognition," in *Proceedings of the 22nd International Society for Music Information Retrieval Conference*, ser. ISMIR 2021, 2021, pp. 443–451.

[70] H. F. Garcia, A. Aguilar, E. Manilow, and B. Pardo, "Leveraging hierarchical structures for few-shot musical instrument recognition," in *Proceedings of the 22nd International Society for Music Information Retrieval Conference*, ser. ISMIR 2021, 2021, pp. 220–228.

[71] X. Hu and J. H. Lee, "A cross-cultural study of music mood perception between American and Chinese listeners," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 535–540.

[72] T. Schäfelr, A. Tipandjan, and P. Sedlmeier, "The functions of music and their relationship to music preference in India and Germany," *International Journal of Psychology*, vol. 47, no. 5, pp. 370–380, 2012.

[73] Y. Yang and X. Hu, "Cross-cultural music mood classification: A comparison on English and Chinese songs," in *Proceedings of the 13th International Society for Music Information Retrieval Conference*, ser. ISMIR 2012, 2012, pp. 19–24.

[74] X. Hu, J. H. Lee, K. Choi, and J. S. Downie, "A cross-cultural study on the mood of K-POP songs," in *Proceedings of the 15th International Society for Music Information Retrieval Conference*, ser. ISMIR 2014, 2014, pp. 385–390.

[75] M. Liu, X. Hu, and M. Schedl, "Artist preferences and cultural, socio-economic distances across countries: A big data perspective," in *Proceedings of the 18th International Society for Music Information Retrieval Conference*, ser. ISMIR 2017, 2017, pp. 103–111.

[76] C. Bauer and M. Schedl, "Global and country-specific mainstreaminess measures: Definitions, analysis, and usage for improving personalized music recommendation systems," *PLOS ONE*, vol. 14, no. 6, pp. 1–36, 2019.