What do Listeners Attend to When Listening to Music? Toward Explainable Music Recommendations

Kosar Seyyedhosseinzadeh[1], Matevž Pesek[2] and Marko Tkalčič[1]

[1]Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia
[2]Faculty of Computer and Information Science, University of Ljubljana, Ljubljana, Slovenia

Abstract

Personalized explanations in recommender systems can be useful when they reflect the aspects of an item that matter to the user. To account for individual preferences regarding different aspects of songs in music recommender systems, it is first necessary to identify which song aspects explain meaningful variation among listeners. This paper introduces a questionnaire instrument designed to operationalize these differences as a measurable user characteristic. The instrument was developed through item generation, expert review, pilot testing, and a main study. Data from the main study ($N = 245$) were analyzed using exploratory factor analysis to examine the questionnaire's internal structure. The results supported a two-factor solution, interpreted as *Lyric-Engagement* and *Music-Engagement*, with both dimensions showing good internal consistency. These findings suggest that listeners' orientations toward lyrical and musical elements can be measured in an interpretable way. The contribution of this study lies not in proposing a new explanation algorithm, but in providing an empirical basis for user modeling that may help align explanation content with the song characteristics most relevant to different listeners.

Keywords

music recommender systems, explainable recommendation, personalized explanations, user modeling, questionnaire development, exploratory factor analysis

1. Introduction

Recommender systems increasingly rely on explanations to improve transparency, trust, and user experience [1, 2]. However, explanations are not universally effective: the same explanation may be meaningful to one user and uninformative to another. A growing body of research suggests that the effectiveness of recommendation explanations depends not only on how the explanations are generated, but also on how users interpret them when evaluating the recommended items [3, 4].

A common implicit assumption in many explanation approaches is that the same item characteristics will be similarly informative for all users [3, 5]. Prior work, however, suggests that users differ in how they perceive and benefit from explanations, and that user characteristics can shape what information is useful to them [5, 4]. In domains such as movies and products, aspect-based recommendation models already assume that items can be described in terms of multiple aspects and that users differ in the aspects they care about [6, 7]. More recent work has continued this line by modeling explainable recommendations in terms of user-specific aspect preferences [8, 9]. Consequently, the effectiveness of an explanation depends in part on whether it is aligned with what is relevant to the user [1].

This perspective suggests that effective explanations should be aligned with the aspects of items that individual users find most relevant. Prior work on personalized explanations has shown that users respond differently to explanation styles depending on their traits, preferences, or cognitive characteristics [3, 4]. The existing literature suggests that personalization is not only a matter of how explanations are presented, but also of which item characteristics are selected and emphasized in the explanation.

Music recommendation provides a particularly suitable domain in which to explore how users respond to explanations grounded in particular aspects of a song [2]. Songs combine multiple components, most

Joint Proceedings of the ACM UMAP Workshops 2026, UMAP 2026, 8–11 June 2026, Gothenburg, Sweden
kosar.seyyedhosseinzadeh@famnit.upr.si (K. Seyyedhosseinzadeh); matevz.pesek@fri.uni-lj.si (M. Pesek); marko.tkalcic@famnit.upr.si (M. Tkalčič)

notably lyrical and musical elements, which play different roles in listeners' experiences. Prior work in music cognition suggests that lyrics and musical structure are not always processed in the same way, and that verbal and musical information can make partly distinct contributions to song perception [10, 11, 12]. This suggests that explanations grounded in different aspects of a song are not equally meaningful for all users.

However, in order to provide explanations tailored to individual preferences for various aspects of songs, one must first operationalize these aspects. In this paper, we propose an instrument for assessing individual differences in listeners' tendency to attend to lyrical and musical aspects of songs. We examine whether these tendencies can be measured in an interpretable manner and suggest that the resulting dimensions provide a useful basis for user modeling in future aspect-based explanation design for music recommender systems. Rather than proposing a new explanation algorithm, this work's contribution lies in identifying which elements of songs are more relevant to a given listener, so that recommendation explanations can be more closely aligned with users' perceptual and evaluative tendencies. The research question addressed in this study is:

Can listeners' orientation toward different aspects of songs, particularly lyrical and musical elements, be operationalized as a measurable construct?

2. Related Work

Here we review related work that supports three main points underlying the present study: first, that recommended items can be understood in terms of multiple characteristics and that users may assign different importance to those characteristics; second, that explanation design and user modeling are closely related when users differ in what they attend to and value; and third, that in music, lyrical and musical elements play distinct roles in listeners' experiences, making them a plausible basis for personalized explanation design.

A central assumption in personalized explainability is that recommended items have multiple characteristics and that users do not assign equal importance to those characteristics. An explanation may therefore be less useful when it emphasizes aspects of an item that are not central to how a particular user evaluates it. This assumption is reflected in several strands of explainable recommendation research, especially in feature-based and aspect-based work. For example, Tintarev and Masthoff [1] showed that personalized feature-based explanations can influence both user satisfaction and decision quality, while also noting that different explanation goals may conflict with one another. In a more explicitly aspect-based line of work, Zhang et al. [6] introduced explicit factor models that make recommendations interpretable in terms of item features learned from reviews, while also accounting for the fact that different users may care about different aspects of the same item. Hou et al. [8] developed this point further by arguing that users may differ in aspect preference even when they assign the same overall rating, and that recommendation and explanation should therefore operate at the aspect level rather than only at the item level. Related work has likewise modeled user preferences through user–item–aspect relations extracted from reviews in order to improve recommendation transparency and explainability [7]. More recent work has extended this line by identifying the aspects users are most concerned about for particular items in order to generate more personalized explanations [9]. Similarly, Pan et al. [13] treated explanation as a matter of linking recommendations to interpretable item characteristics rather than leaving them at the level of latent factors.

A second line of research supports the idea that explanation design should take differences in user priorities into account. Work on critiquing-based recommendation has treated preference construction as an iterative process in which users indicate which attributes matter more or less to them, and the system updates its recommendations accordingly [14]. Interactive systems such as TasteWeights made such differences even more explicit by allowing users to inspect and adjust recommendation weights directly [15]. Related work has also shown that recommendation interfaces can be designed around the assumption that users benefit from seeing and influencing the factors behind recommendations [16]. In hybrid recommenders, explanation usefulness has likewise been shown to depend on both user characteristics and explanation format [3]. Overall, this literature suggests that personalization is not only a matter of presentation style, but also of matching explanation content to what is likely to matter to the user.

Music recommendation poses a more specific version of the broader problem outlined above: namely, which elements of a song should serve as the basis of an explanation when listeners may assign different degrees of importance to them. This is especially relevant in music because songs are short, highly context-dependent, and represented through multiple types of information, including audio, lyrics, metadata, and interaction behavior. As noted in a review of explainability in music recommender systems, explanation methods cannot simply be transferred from other domains without adaptation [2]. This problem is not only a design issue for music recommender systems, but also a theoretically plausible one, because research in music psychology suggests that lyrical and musical elements may be processed differently and may contribute differently to listeners' experiences. For example, Besson et al. [10] found that semantic anomalies in lyrics and tonal anomalies in melody elicited different neural responses, indicating that verbal and musical material are not processed in the same way. Bonnel et al. [11] likewise showed that listeners can allocate attention differently to lyrics and tunes during song processing. Other studies found that lyrics and melody can contribute differently to emotional response [17], and that musical valence can influence the cognitive processing of lyrics [12]. A broader perspective is provided by the Goldsmiths Musical Sophistication Index, which treats musical engagement as a multidimensional construct involving listening habits, perceptual orientation, and emotional engagement rather than formal training alone [18]. Collectively, these studies suggest that listeners may differ not only in their musical experience, but also in the relative importance they assign to lyrical and musical aspects of songs.

What this literature does not establish, however, is whether differences in attention to lyrical and musical elements can be formalized as a measurable user characteristic that could support explanation design in music recommender systems. The present study addresses this gap by asking whether listeners' orientation toward lyrical, as opposed to musical, elements can be measured and interpreted in a meaningful way through a questionnaire instrument.

3. Method

The method was designed to answer the research question by developing a questionnaire intended to operationalize listeners' orientation toward different aspects of songs, particularly lyrical and musical elements, as a measurable construct. The study proceeded in two parts. First, the questionnaire was developed and refined through an exploratory, multi-stage process that included item candidate generation, expert review, pilot testing, and the main study. Second, data from the main study were analyzed to examine the measurement properties of the instrument, with particular attention to its internal structure and psychometric performance.

3.1. Questionnaire Development

Questionnaire development followed a staged process consisting of item candidate generation, expert review, pilot testing, and the main study. The questionnaire developed for the item candidate generation stage was designed to cover a broad range of potentially relevant aspects of how listeners attend to different elements of songs. At this stage, the aim was not to impose a predefined factor structure, but rather to generate a broad set of items that could later be evaluated and refined. Before pilot testing, the full item pool was reviewed by an expert for clarity, relevance, and redundancy, and revised accordingly.

During the pilot testing phase, the questionnaire was administered online in a self-paced format using LimeSurvey, an online survey platform for creating and distributing questionnaires. After completing the information and consent page, participants answered two groups of questions. The first group consisted of 31 core items intended to measure the construct, each rated on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The second group included 12 background questions on listening habits, music education, and demographic characteristics. Some of the music-related

Table 1
Demographic and music-related characteristics of the final analytical sample

Measure	Value
Total Participants	245
Age	18–79 years ($M = 35.39$, $SD = 12.68$)
Gender	42.9% female; 55.5% male; 1.2% non-binary / another gender identity; 0.4% prefer not to say
Region of origin	Europe (60.4%); America (30.2%); Asia (4.1%); Africa (2.4%); Middle East (1.2%); Other (1.6%)
Music training	33.1% none; 15.9% informal training; 26.9% school music classes; 18.0% private lessons; 3.3% conservatory / pre-professional training
Daily music listening	1–2 hours (39.1%); less than 1 hour (20.0%); 2–3 hours (19.5%)
Most frequently listened genres	Pop (42.3%), Rock (36.2%), and Hip-Hop / Rap (30.1%)
Least frequently listened genres	Reggae / Dub / Ska (12.6%), Latin / Afro-Caribbean (15.9%), and Jazz (16.4%)
Listening context	67.8% usually listened alone; 31.8% reported that it depended on the situation

background questions were informed by the broader concept of musical sophistication reflected in the Gold-MSI [18]. Four attention-check items were distributed throughout the survey. Participants were recruited from a university student population, and 45 completed the questionnaire. Participants could also provide optional written comments at the end of the survey. The median completion time was approximately 9 minutes. Revisions after this stage were based on both participant feedback and item-level psychometric evaluation, including corrected item-total correlations and internal consistency estimates derived from Cronbach's alpha and McDonald's omega. Items showing weak performance or redundancy were revised or removed.

The main study was then administered online to a larger sample recruited through Prolific, a web-based participant recruitment platform commonly used in behavioral and survey research. The questionnaire contained 28 core items intended to measure the construct, 4 attention-check items, and 13 background questions, including one additional question compared with the pilot testing version on how much time participants typically spent listening to music each day. The data collected in the main study were used for the exploratory factor analysis reported in this paper. As part of the recruitment criteria, only Prolific users with an approval rate above 95% were eligible to participate. Participants first viewed an information and consent page, then completed the core items, followed by the background questions, and finally received a completion code through Prolific after submitting the questionnaire. The median completion time at this stage was approximately 7 minutes. A total of 250 respondents completed the questionnaire. Five respondents failed the attention checks and were excluded, resulting in a final analytical sample of $N = 245$. The demographic and music-related characteristics of this final analytical sample are summarized in Table 1. The data from this sample were then used for the exploratory psychometric analyses described in the next subsection.

3.2. Data Analysis

The analyses were conducted in two stages, corresponding to the pilot testing and the main study.

In the pilot testing stage, the item sets developed during the initial item generation stage were examined with the aim of refining and revising the questionnaire. To this end, after collecting the pilot data, descriptive statistics were inspected for each item set, including response distributions, means, and variability, in order to identify items with limited variance or problematic response patterns. Internal consistency was assessed using Cronbach's α and McDonald's ω, which provide estimates of how coherently items operate together within a provisional item set, with α based on average inter-item correlations and ω offering a less restrictive estimate that does not assume equal item loadings. These results were used to guide item revision, removal, and shortening of the questionnaire, with particular attention to items showing low consistency, redundancy, or unclear interpretation.

In the main study, exploratory factor analysis (EFA) was used to examine the questionnaire's underlying dimensions and initial psychometric properties. Because the study concerned early-stage questionnaire development, no factor structure was imposed in advance, and all analyses were conducted exploratorily. For the main study data, EFA was used to identify the underlying dimensions of the questionnaire, as the latent structure of the item pool was not known in advance. Before conducting the EFA, the suitability of the data for factor analysis was assessed using Bartlett's test of sphericity and the Kaiser–Meyer–Olkin measure of sampling adequacy. Factors were extracted using minimum residual estimation. For solutions with more than one factor, oblimin rotation was applied because the underlying listening dimensions were expected to be related rather than fully independent. Factor retention was evaluated using eigenvalues, scree inspection, and parallel analysis, as these criteria provide complementary evidence about the number of factors to retain, with parallel analysis offering a stronger empirical benchmark than eigenvalues alone. Candidate solutions were then compared with respect to interpretability, primary loadings, cross-loadings, and communalities to identify the most coherent and interpretable item structure.

4. Results

The results are organized in two parts. First, exploratory factor analysis was used to examine whether the questionnaire showed an interpretable internal structure. Second, the reliability of the retained dimensions was assessed.

4.1. Exploratory Factor Analysis

Exploratory factor analysis (EFA) was conducted on the 28 core items intended to measure the construct in the main study questionnaire, using data from the final analytical sample ($N = 245$). Before analysis, negatively worded items were reverse-coded, and attention-check items were excluded. Preliminary checks indicated that the data were suitable for factor analysis. Bartlett's test of sphericity was significant, $\chi^2 = 2788.27, p < .001$, and the Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy indicated very good overall factorability (KMO = .897). Item-level KMO values were also consistently high, ranging from .809 to .943.

Initial eigenvalues suggested several possible factors, with the first six eigenvalues exceeding 1.0. However, parallel analysis provided clearer evidence for a two-factor solution: only the first two observed eigenvalues exceeded both the mean and the 95th percentile of the corresponding random-data eigenvalues. On this basis, one- through six-factor solutions were estimated and compared.

The one-factor solution accounted for 26.5% of the variance, but was considered too restrictive to capture the structure of the item pool. The two-factor solution accounted for 35.4% of the variance. It yielded the clearest and most interpretable pattern of loadings, with 23 items clearly loading, relatively few cross-loadings, and fewer weak items than the higher-factor solutions. Although the three-, four-, five-, and six-factor solutions explained slightly more variance, they also introduced a larger number of weakly loading or conceptually diffuse items, reducing interpretability.

Inspection of the two-factor loading pattern suggested that the retained structure reflected two broader factors, interpreted here as distinct listening orientations. The first factor was defined by items related to attention to musical aspects of songs and was labeled *Music-Engagement*. The second factor was defined by items related to attention to lyrics and was labeled *Lyric-Engagement*. Applying a primary-loading threshold of $\geq .40$ and allowing only limited cross-loading resulted in the retention of 23 items in total, comprising 12 music-engagement items and 11 lyric-engagement items. The retained questionnaire items and their loadings on the *Lyric-Engagement* and *Music-Engagement* factors are presented in Table 2. The five excluded items are presented in Table 3 and showed high cross-loading, and their wording was also less clearly aligned with only one of the two dimensions. These items were therefore not treated as clearly belonging to either the *Lyric-Engagement* or *Music-Engagement* factor and were excluded to improve the interpretability of the final two-factor solution.

Table 2
Retained items from the final two-factor solution. The inter-factor correlation was $r = .29$.

Question	Lyric-Engagement	Music-Engagement
Song lyrics are an important part of my listening experience.	.851	-.136
The quality of a song's lyrics matters a lot to me.	.756	-.041
I often think about a song's lyrics after listening to it.	.674	.087
Song lyrics are not something I usually use to express my emotions.	.655	.002
I don't usually memorize the lyrics of my favorite songs.	.576	.045
Song lyrics do not usually stay in my mind.	.571	-.007
I appreciate lyrics that are expressive, clever, or meaningful.	.549	.079
I actively form a narrative from a song's lyrics while listening.	.536	.240
The lyrics I connect with are not typically related to what I believe or who I am.	.494	-.144
When a singer starts singing, I shift my focus to their voice.	.438	.264
When I hear a beautiful snippet of a song's lyrics, I want to listen to the full song.	.417	.116
I do not usually focus on the sound of individual instruments in a song.	-.160	.708
Instrumental sections of songs frequently capture my attention.	-.230	.694
I notice when instruments repeat patterns or change them.	-.116	.686
I'm drawn to a song that emphasizes musicianship and composition.	.034	.644
The musical qualities of a song are an important part of my listening experience.	-.024	.639
I do not usually notice when more or fewer musical parts are playing in a song.	.011	.593
I do not usually pay attention to whether a singer's voice is in tune and in time.	.120	.590
I tend to notice when a singer's voice gently goes up and down on a single note.	.238	.560
I do not usually notice when a singer switches to a lighter or heavier vocal tone.	.101	.541
I pay close attention to a singer's vocal technique.	.241	.510
How the lead and backing singers blend together does not usually draw my attention.	.081	.465
I can recognize a familiar singer from a short clip of a song.	.175	.443

These findings suggest that the questionnaire is best represented by two related listening orientations rather than by a single undifferentiated construct or a more fragmented multidimensional structure. As shown in Figure 1, only the first two observed eigenvalues exceeded the corresponding random-data benchmarks.

4.2. Reliability of the Retained Dimensions

Internal consistency was assessed using Cronbach's alpha (α) and McDonald's omega (ω) for the retained item sets derived from the two-factor solution.

As summarized in Table 4, the *Lyric-Engagement* dimension, consisting of 11 items, showed good internal consistency ($\alpha = .862$, $\omega = .868$). The *Music-Engagement* dimension, consisting of 12 items, also demonstrated good reliability ($\alpha = .869$, $\omega = .872$).

These results indicate that both retained dimensions were measured with good reliability in the final

Table 3
Items excluded from the final two-factor solution.

Question	Lyric-Engagement	Music-Engagement
A singer's voice often stays in my mind for a long time.	.422	.329
I pay close attention to how the lyrics fit the rhythm and phrasing of the music.	.365	.351
I notice when a singer adjusts their volume intentionally for emotional effect	.310	.494
The different sections of a song (e.g., verses, choruses, bridges) do not typically draw my attention.	.090	.363
I enjoy how different singers bring their own interpretation to the same song.	.224	.283

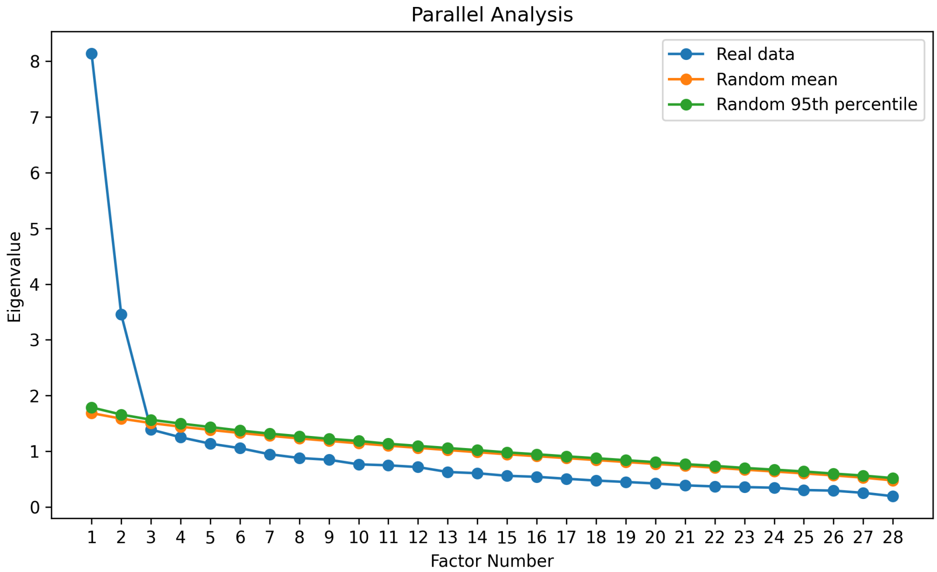

Figure 1: Parallel analysis results for factor retention. Only the first two observed eigenvalues exceeded the corresponding random-data benchmarks, supporting a two-factor solution.

Table 4
Summary of the retained two-factor solution and reliability estimates

Factor	Interpretation	Items	Variance	α	ω
F1	Music-Engagement	12	18.4%	.869	.872
F2	Lyric-Engagement	11	17.0%	.862	.868

analytical sample. The two retained factors were moderately correlated ($r = .29$), supporting the use of an oblique rotation; the reliability estimates support treating them as distinguishable but related listening orientations.

5. Discussion

The two dimensions identified in this study provide a basis for personalizing which characteristics of a song can be selected as explanation content. For listeners whose orientation is more strongly lyrical, explanations may be more meaningful when they emphasize lyrical content and verbal meaning. For listeners whose orientation is more strongly musical, explanations may be more useful when they focus on musical properties such as melody, rhythm, or harmony. Rather than pointing to clearly separate listener types, the results indicate continuous variation along these two dimensions. The contribution of the present study is therefore not a new explanation algorithm, but a user-modeling step that helps determine which aspects of a song should ground an explanation for a given listener. In this way, the proposed instrument supports a more personalized alignment between explanation content and listeners' perceptual and evaluative tendencies.

Several limitations should be acknowledged. First, the analysis is exploratory, and the factor structure should therefore be examined further in independent samples and with confirmatory methods. Second, the data are based on self-reports and thus reflect perceived listening orientation rather than observed listening behavior. Third, although the study is motivated by explainability in recommender systems, no recommender interface was implemented, and no user-facing evaluation of explanation strategies was conducted.

Overall, the study supports treating listeners' orientation toward lyrical and musical elements as a measurable user characteristic that can inform explanation design in music recommender systems.

Acknowledgments

This work was supported by the ARIS funding agency through the grant number N2-0354.

Declaration on Generative AI

During the preparation of this work, the authors used ChatGPT for language proofreading and writing assistance. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] N. Tintarev, J. Masthoff, Evaluating the effectiveness of explanations for recommender systems: Methodological issues and empirical studies on the impact of personalization, User Modeling and User-Adapted Interaction 22 (2012) 399–439. URL: http://link.springer.com/10.1007/s11257-011-9117-5. doi:10.1007/s11257-011-9117-5.

[2] D. Afchar, A. B. Melchiorre, M. Schedl, R. Hennequin, E. V. Epure, M. Moussallam, Explainability in music recommender systems, AI Magazine 43 (2022) 190–208. URL: https://onlinelibrary.wiley.com/doi/10.1002/aaai.12056. doi:10.1002/aaai.12056.

[3] P. Kouki, J. Schaffer, J. Pujara, J. O'Donovan, L. Getoor, Personalized explanations for hybrid recommender systems, in: Proceedings of the 24th International Conference on Intelligent User Interfaces, ACM, Marina del Ray, California, 2019, pp. 379–390. URL: https://dl.acm.org/doi/10.1145/3301275.3302306. doi:10.1145/3301275.3302306.

[4] M. Millecamp, C. Conati, K. Verbert, "Knowing me, knowing you": personalized explanations for a music recommender system, User Modeling and User-Adapted Interaction 32 (2022) 215–252. URL: https://link.springer.com/10.1007/s11257-021-09304-9. doi:10.1007/s11257-021-09304-9.

[5] K. Wardatzky, O. Inel, L. Rossetto, A. Bernstein, Whom do Explanations Serve? A Systematic Literature Survey of User Characteristics in Explainable Recommender Systems Evaluation, ACM Transactions on Recommender Systems 3 (2025) 1–35. URL: https://dl.acm.org/doi/10.1145/3716394. doi:10.1145/3716394.

[6] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, S. Ma, Explicit factor models for explainable recommendation based on phrase-level sentiment analysis, in: Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval, ACM, Gold Coast, Queensland, Australia, 2014, pp. 83–92. URL: https://dl.acm.org/doi/10.1145/2600428.2609579. doi:10.1145/2600428.2609579.

[7] X. He, T. Chen, M.-Y. Kan, X. Chen, TriRank: Review-aware Explainable Recommendation by Modeling Aspects, in: Proceedings of the 24th ACM International Conference on Information and Knowledge Management, ACM, Melbourne, Australia, 2015, pp. 1661–1670. URL: https://dl.acm.org/doi/10.1145/2806416.2806504. doi:10.1145/2806416.2806504.

[8] Y. Hou, N. Yang, Y. Wu, P. S. Yu, Explainable recommendation with fusion of aspect information, World Wide Web 22 (2019) 221–240. URL: https://doi.org/10.1007/s11280-018-0558-1. doi:10.1007/s11280-018-0558-1.

[9] H. Cheng, S. Wang, W. Lu, W. Zhang, M. Zhou, K. Lu, H. Liao, Explainable Recommendation with Personalized Review Retrieval and Aspect Learning, in: A. Rogers, J. Boyd-Graber, N. Okazaki (Eds.), Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Toronto, Canada, 2023, pp. 51–64. URL: https://aclanthology.org/2023.acl-long.4/. doi:10.18653/v1/2023.acl-long.4.

[10] M. Besson, F. Faïta, I. Peretz, A.-M. Bonnel, J. Requin, Singing in the Brain: Independence of Lyrics and Tunes, Psychological Science 9 (1998) 494–498. URL: https://doi.org/10.1111/1467-9280.00091. doi:10.1111/1467-9280.00091.

[11] A.-M. Bonnel, F. Faita, I. Peretz, M. Besson, Divided attention between lyrics and tunes of operatic songs: Evidence for independent processing, Perception & Psychophysics 63 (2001) 1201–1213. URL: http://link.springer.com/10.3758/BF03194534. doi:10.3758/BF03194534.

[12] A. Fiveash, G. Luck, Effects of musical valence on the cognitive processing of lyrics, Psychology of Music 44 (2016) 1346–1360. URL: https://journals.sagepub.com/doi/10.1177/0305735615628057. doi:10.1177/0305735615628057.

[13] D. Pan, X. Li, X. Li, D. Zhu, Explainable Recommendation via Interpretable Feature Mapping and Evaluation of Explainability, in: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, International Joint Conferences on Artificial Intelligence Organization, Yokohama, Japan, 2020, pp. 2690–2696. URL: https://www.ijcai.org/proceedings/2020/373. doi:10.24963/ijcai.2020/373.

[14] L. Chen, P. Pu, Critiquing-based recommenders: survey and emerging trends, User Modeling and User-Adapted Interaction 22 (2012) 125–150. URL: http://link.springer.com/10.1007/s11257-011-9108-6. doi:10.1007/s11257-011-9108-6.

[15] S. Bostandjiev, J. O'Donovan, T. Höllerer, TasteWeights: a visual interactive hybrid recommender system, in: Proceedings of the sixth ACM conference on Recommender systems, ACM, Dublin, Ireland, 2012, pp. 35–42. URL: https://dl.acm.org/doi/10.1145/2365952.2365964. doi:10.1145/2365952.2365964.

[16] B. P. Knijnenburg, S. Bostandjiev, J. O'Donovan, A. Kobsa, Inspectability and control in social recommenders, in: Proceedings of the sixth ACM conference on Recommender systems, ACM, Dublin Ireland, 2012, pp. 43–50. URL: https://dl.acm.org/doi/10.1145/2365952.2365966. doi:10.1145/2365952.2365966.

[17] S. O. Ali, Z. F. Peynircioğlu, Songs and emotions: are lyrics and melodies equal partners?, Psychology of Music 34 (2006) 511–534. URL: https://journals.sagepub.com/doi/10.1177/0305735606067168. doi:10.1177/0305735606067168.

[18] D. Müllensiefen, B. Gingras, J. Musil, L. Stewart, The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population, PLOS ONE 9 (2014) e89642. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0089642. doi:10.1371/journal.pone.0089642.