Towards Personality-Aware Explanations for Music Recommendations Using Generative AI

Gabrielle Alves*
Universidade de São Paulo
São Carlos, Brazil
gaah.alves@usp.br

Dietmar Jannach
University of Klagenfurt
Klagenfurt, Austria
dietmar.jannach@aau.at

Luan Soares de Souza*
Universidade de São Paulo
São Carlos, Brazil
luanssouza@usp.br

Marcelo Garcia Manzato*
Universidade de São Paulo
São Carlos, Brazil
mmanzato@icmc.usp.br

Abstract

It is well established that the provision of explanations can positively impact the effectiveness of a recommender system. In many proposals in the literature, these explanations are personalized in that they refer to a user's known individual preferences. Some recent works, however, also indicate that personalization should also happen at a higher level, where the system, in a first step, decides in which specific way an explanation should be provided, depending, for example, on the user's expertise. In this research, we take the first steps towards personality-aware explanations by exploring how users perceive explanations tailored to reflect the Big Five personality traits. To this purpose, we leverage the capabilities of modern Generative AI tools to create personality-based explanations at scale in the context of a music recommendation scenario. A linguistic analysis of the generated explanations confirms that they properly reflect expected language patterns associated with individual personality traits. Furthermore, a user study shows that users tend to prefer certain linguistic framings over others, for example, explanations that reflect low-neuroticism patterns. In addition, we find that some explanation forms are more effective than others regarding persuasiveness and perceived overall quality.

CCS Concepts

• **Information systems** → **Recommender systems**.

Keywords

Recommender Systems, Explanations, Personality, LLM

ACM Reference Format:
Gabrielle Alves, Dietmar Jannach, Luan Soares de Souza, and Marcelo Garcia Manzato. 2025. Towards Personality-Aware Explanations for Music Recommendations Using Generative AI. In *Proceedings of the Nineteenth ACM Conference on Recommender Systems (RecSys '25), September 22–26, 2025, Prague, Czech Republic.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3705328.3748032

*Supported by CNPq, CAPES, FAPESP, and INCT-TILD-IAR, process number 408490/2024-1.

ACM ISBN 979-8-4007-1364-4/25/09
https://doi.org/10.1145/3705328.3748032

1 Introduction

Recommender systems can benefit in various ways from being able to explain their suggestions to users [21, 42, 46]. Accordingly, a number of explanation approaches have been put forward in the literature [15, 34, 36]. Many of these approaches are based on personalizing the provided explanatory information [41]. For example, instead of informing the user that a given item is popular in general, various approaches relate user preferences with item features (e.g., *"because you like action movies"*) [4, 8, 44].

Recent works indicate that it can also be beneficial to apply personalization at a higher level [2], because not every way of explaining a recommendation may be equally suited for every user. Some users might prefer more detailed explanations, while others like concise explanations [9, 18]. Similarly, some users might prefer textual explanations over visual ones and vice versa [25]. An influential study by Millecamp et al. [30] actually revealed that personal characteristics like *need for cognition* or *domain expertise* may impact the perceived benefit of explanations. Furthermore, in a recent work, Fatahi et al. [13] found that people with different *personality traits* are receptive to different persuasive explanatory messages provided by the system.

In the work by Fatahi et al. [13], the user perception of a predefined set of persuasive messages was analyzed. These messages implement established persuasive strategies from the literature [10]. In our present work, we wonder how users would perceive explanations if they were written in the linguistic style associated with specific personality traits. For example, an explanation designed to match characteristics of high Openness might highlight unique or imaginative aspects of a recommended item. In contrast, an explanation reflecting Conscientiousness might adopt a more structured tone and emphasize technical or practical details. Rather than tailoring explanations to individual users' personalities, we focus on how people respond to different trait-aligned framings of the recommendation.

The long-term goal of our research is to develop design guidelines for explanations that consider the personality traits of individual users, analogously to what has been done in [13]. As a first step towards this goal, we investigate to what extent modern Generative AI technology—in our case in the form of ChatGPT—allows us to generate textual explanations for different personality traits at scale. As an application use case, we focus on music recommendation, an important application area of personality-aware systems [12]. To study ChatGPT's capabilities to generate personality-trait aware explanations, we engineered appropriate prompts and performed a linguistic analysis, which confirmed that the explanations matched the linguistic patterns that one would expect according to the Big Five personality model [17]. Furthermore, we conducted a user study involving N=348 users to gauge the perception of such explanations. The study revealed that not all types of explanations are equally preferred by the users and that the quality assessment of the explanations also varied. As a result, our study provides us with the first important insights regarding our longer-term goal of personality-aware explanations for recommender systems.

The paper is organized as follows. We discuss previous work next in Section 2. Section 3 describes our methodology, and in Section 4 we present our results. The paper closes with a discussion of research limitations and future work.

2 Previous Work

The literature on explainable recommendations is generally rich, and we refer readers to existing surveys in this area [21, 42, 46]. Here, we highlight selected recent studies that are closely related to our work.

Silva et al. [39] used ChatGPT to generate both recommendations and corresponding explanations for users, and they compare personalized and generic explanations in terms of perceived personalization, persuasiveness, and effectiveness. While personalized recommendations led to higher user satisfaction, personalizing the explanations of the recommended items had no measurable effect on the studied user perceptions, including the perceived level of personalization. A related observation was made earlier by Balog and Radlinski [5], who found that 'neutral' (non-personalized) explanations *"performed particularly well"* in their study that aimed to investigate quality perceptions of different forms of goal-oriented explanations. While our work also focuses on user perceptions of different types of explanations, we differ from [5, 39] in that our goal is to investigate the feasibility and perception of personality-aware explanations.

The work by Matz et al. [28] is very close to ours, as they also explore the capability of ChatGPT to generate psychologically-tailored messages to users at scale based on the Big Five personality framework. Overall, several studies found that the generated personalized messages exhibited more influence on users than non-personalized ones. The focus of their work, however, is on persuasion—encouraging behaviors in domains such as marketing, health, and politics, whereas we study the perception of explanations in the context of recommender systems. Furthermore, unlike our work, Matz et al. [28] did not perform a lexical analysis to verify whether the generated content reflects personality-specific linguistic patterns.

Regarding further related works, Li et al. [27] also rely on Generative AI, in their case building on a GPT-based model, to create explanations in natural language in an automated way based on prompt learning. Unlike our work, however, the generated explanations mostly revolve around item features and are not tailored to be personality-aware. Fatahi et al. [13], as mentioned above, assess the effectiveness of a fixed set of persuasion strategies depending on user personalities. Like our study, they rely on the music domain and the Big Five model, but the focus, similarly to [28], is on persuasion (see also [3, 13, 16, 23]), and differently from our work they do not rely on Generative AI to create personality-aware explanations. In a very recent work, Noughabi et al. [33] follow similar ideas as [13], exploring the effectiveness of predefined personality-based explanations with the goal of persuasion.

The work by Millecamp et al. [30] also targets the music domain, and they study how personal characteristics may affect the perception of explanations. Their user study concluded that these characteristics should be considered when designing explanations. Instead of personality traits, they consider different personal factors such as musical experience or visualization literacy. The primary focus of the work by Berkovsky et al. [7] is how content selection for explanations impacts the users' trust in the system, but they also study whether trust depends on user personality traits. However, the consideration of personality traits is quite different from our work, where our goal is to explore ways to make the explanation content itself personality-aware.

Finally, we note that our user study concentrates on recommendations in the music domain. In this area, the user personality traits seem particularly important [12] and are correlated with the users' preferred genres or song features [6, 14, 25, 29].

3 Methodology

We recall the two main research questions that we seek to answer in this first step towards enabling personality-aware explanations in recommender systems:

- RQ1: Can ChatGPT generate explanations that reflect the expected language patterns associated with specific personality traits when prompted accordingly?
- RQ2: How do users perceive explanations of different types in dimensions such as quality and persuasiveness?

3.1 Generating and Assessing Explanations

Prompt Engineering Approach. We followed an iterative approach to prompt engineering, considering the outputs of multiple LLMs (ChatGPT, Gemini, LLaMA) during the process[1] to evaluate their ability to generate distinct explanations. Ultimately, we relied on ChatGPT-4o, which was prompted to create five alternative explanations for a given song[2], each matching one of the five psychological traits of the Big Five model [22] (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism). For each given trait, a prompt was created that described behavioral characteristics associated with the trait (e.g., "talkative", "assertive" for extroversion). Furthermore, the prompt instructed ChatGPT to exclude the song title and artist name and to avoid directly mentioning the personality terms in the explanation. The prompt for the extroversion-oriented explanation reads as follows[3].

[1]The model outputs were manually verified to ensure linguistic plausibility, internal coherence, and adherence to framing constraints.

[2]The set of available songs is determined by the dataset that we used in the subsequent user study.

[3]Translated to English as the study was conducted in Portuguese; the word limit was chosen to avoid cognitive overload during the user study.

"Write an explanation of why <Song> by <Artist> would be a good choice for someone who is extroverted and enthusiastic (that is, sociable, assertive, talkative, active, NOT reserved or shy). Use exactly 50 words, do not mention the song's name or artists. Explain why the sound is perfect for these personality traits without explicitly mentioning words like extroverted, enthusiastic, sociable, assertive, talkative, active."

In addition to the five personality-based explanations, a feature-based prompt and explanation served as the control condition, focusing on non-emotional song attributes such as rhythm, structure, and instrumentation. To support reproducibility, all exact prompts are publicly available in the online material.[4]

Assessment Approach. We performed a linguistic and statistical analysis to determine if the central words in the generated explanations correspond to concepts that one would find in the psychology literature to describe a given personality trait [20, 26, 43]. For this analysis, we first removed stop words from the text. Then, we analyzed the word frequencies and found that a relatively small fraction of the terms accounts for most of the observed term frequencies. Following the Pareto principle, we focused on the top 20% of terms by frequency and omitted the rest. This cleaning process also contributes to the stability of the subsequent statistical tests.

Following the approach by Oakes and Farrow [35], we then applied a Chi-Squared test of independence to examine the association between specific words in the explanations and the given personality traits. This test allowed us to determine whether the distribution of words across trait-specific explanation sets differed significantly from what would be expected. Standardized residuals were calculated to interpret the strength and direction of these associations. To assess the lexical alignment with the psychology literature, we developed an enriched trait lexicon by combining prior word lists with an expanded set based on Goldberg's Big Five markers [17] and supplemented it with terms adapted from the psychology literature [22, 26, 40, 43]. Each top word was matched to a trait if its stemmed form aligned with a lexicon term, and Precision, Recall, and F1-scores were calculated to quantify the overall alignment with established psychological descriptors[5].

3.2 Assessing User Perceptions

To study user preferences and perceptions (RQ2), we designed a within-subjects online experiment in which each participant was shown six music recommendations in the form of ChatGPT-based explanations. The six conditions consisted of five personality-based explanations, each generated for a different trait, plus the explanation solely based on song features. To answer RQ2, the participants were instructed (a) to select their most preferred recommendation/explanation (*Preference*), a choice which directly determined the song they would listen to on the following screen, thus allowing us to observe actual user behavior, and (b) to answer a set of 5-point Likert-scale questions regarding their preferences and behavioral intentions. Specifically, for task (b), participants were asked for each recommendation (i) if they would like to add the song to their playlist (*Persuasiveness*), (ii) how they rate the quality of the recommendation *(Quality)*, and (iii) to what extent the recommendation matches their musical taste *(Accuracy)*[6].

Regarding the overall study workflow, participants first provided their informed consent, which was developed in compliance with the relevant national regulations. Then, they were asked to select at least seven songs they liked from a pre-filtered list of songs released between 2000 and 2020. The collaborative filtering algorithm SLIM [32] was then used to provide personalized recommendations. These recommendations were shown as generated explanations (without disclosing song or artist names) as described above. The item-based SLIM model allowed us to generate personalized recommendations for new users without model retraining. The explanations were generated on the fly through OpenAI's API. The order of the explanations was randomized for each participant. Screen captures of the user interface are provided in the online material, along with a description of the song dataset used.

4 Results

In this section, we present and discuss the outcomes of our analyses.

4.1 RQ1: Feasibility of Personality-aware Explanations At Scale

We processed 2,646 generated explanations as described in Section 3.1, leading to a set of 188 most relevant terms. We recall that we used a chi-squared (χ^2) test to determine whether the observed frequencies significantly varied across different personality-based explanation types. The test revealed a high χ^2 statistic (50978.37, $p < 0.0001$, df = 935), suggesting that certain words were more strongly associated with specific traits than expected by chance. Cramér's V (0.43) suggests a moderate to strong association. Based on the standardized residuals, we then identified those terms that strongly contribute to the observed χ^2 statistic, and thus are the most distinctive and representative of each personality trait[7]. For each explanation type, approximately 35 to 40 words showed a significant association. We provide all the lists of relevant terms and the residuals in the online material.

Comparing these terms with how the different traits are described in the psychology literature, as described in Section 3.1, we find a strong alignment. For Openness, words such as *exploration, stimulate, imagination, and innovative* reflect novelty-seeking and cognitive flexibility, consistent with descriptors like curious, imaginative, and adventurous, achieving the highest F1-score within the Openness trait words (F1 = 0.269). For Conscientiousness, terms like *focus, precision, and planning* align with established markers of organization and self-discipline (F1 = 0.137). Agreeableness was reflected in emotionally warm terms like *empathy, genuine, and understanding* (F1 = 0.152). Extroversion was characterized by high-energy social terms such as *interaction, vibrant, and movement* (F1 = 0.111). Notably, Neuroticism showed a divergence. The most important terms, including *help, relief* and *comfort*, reflected emotional stability rather than distress, achieving the highest F1-score within

[4]https://github.com/personalities-and-explanations/RecSys25

[5]Top words were translated into English when necessary and stemmed using the Porter stemmer to standardize word forms.

[6]The exact questions are provided in the online material. To not overload the study participants, we limited ourselves to three central questions regarding recommendation and explanation quality. For more comprehensive evaluation frameworks, we refer readers to [24, 37].

[7]Residuals greater than 4 were considered to be strong contributors [1].

the Low Neuroticism trait words (F1 = 0.286). Apparently, ChatGPT in these cases targeted *low* neuroticism, while in other cases it personalized the explanations for the *high* condition of the trait. We attribute this behavior to ChatGPT's general tendency to avoid language associated with negative affect.

Overall, our analysis provides solid empirical evidence suggesting that generative AI approaches like ChatGPT can create personality-aware explanations at scale (RQ1).

4.2 RQ2: User Preferences and Perceptions

We recruited N=348 study participants[8] by inviting followers of a social book review site. Most participants (87%) were female, and 73% were between 26 and 35 years old. On average, they took 15 minutes to complete the study.

Preference. As described in Section 3.2, study participants were asked to select their single most preferred explanation. Figure 1 shows the observed distribution. A χ^2 goodness-of-fit test indicated that the choices were not evenly distributed (p<0.001). Follow-up pairwise χ^2 tests with Bonferroni correction revealed several significant differences. Feature-based explanations were preferred over ones that target conscientiousness, openness and extroversion. Conscientiousness-oriented explanations were less preferred than those focusing on agreeableness and neuroticism.

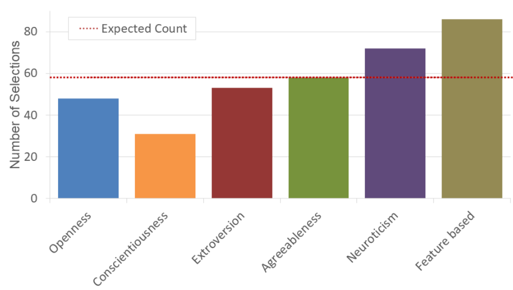

Figure 1: Distribution of user preference across explanation types.

Given these observations, we can confirm that how recommendations are explained in terms of user preference makes a difference. While neutral, feature-based explanations were preferred most frequently, other forms of explaining were found compelling by many participants as well. The relatively strong preference for low neuroticism and weak preference for high conscientiousness seem 'natural' for the music domain, where music is often used as a means to steer emotions (neuroticism) [31, 38] and musical choice is often not a highly planned activity (conscientiousness).

Persuasiveness. For each recommendation, participants rated how likely they were to add the song to the playlist. The average responses for Persuasiveness, Quality, and Accuracy are shown in Table 1. A repeated-measures Friedman test revealed a significant difference in ratings across explanation types ($\chi^2(5) = 45.05$, p<.001). Bonferroni-corrected post-hoc Wilcoxon tests point to a number of significant pairwise differences. In line with the findings regarding *Preference*, it was found that the conscientiousness-oriented explanation was significantly less persuasive than all other explanation forms except for the openness-oriented explanations, which were also found to be of limited persuasiveness.

An example for a conscientiousness-based explanation reads as follows: *"With energetic beats and encouraging lyrics, the melody inspires perseverance and resilience, promoting focus during difficult moments. The electronic harmony is meticulously structured, reflecting precision and care [...]"*. Intuitively, an explanation like this may only have persuasive potential when the listener is in a particular situation requiring focus and attention.

Quality. The same statistical tests as for *Persuasiveness* were applied also for the participants' quality rating—on a scale from 1 to 5—of the provided explanations. Statistically significant differences were observed when considering all groups (p<0.001). Like for *Persuasiveness*, the conscientiousness-oriented explanations stood out in a negative sense. Post-hoc tests revealed that the quality rating for these explanations was significantly lower than for the other explanations, except for openness-oriented explanations.

DV/Construct	Open.	Cons.	Extra.	Agree.	Neuro.	Feature
Persuasiveness	3.89	3.82	4.01	4.03	4.13	4.07
Quality	4.23	4.14	4.32	4.34	4.33	4.35
Accuracy	3.95	3.91	4.03	4.09	4.16	4.15

Table 1: Mean scores for each explanation type.

Accuracy. Regarding whether the provided recommendation/explanation matched the user's general taste, we again found significant differences between the groups (p<0.001). Like the other measurements regarding the users' perception, the explanations tailored for conscientiousness and openness stood out by often being considered to be of limited match with past user preferences. Notably, the differences between the conscientiousness-focused explanations were found to be significantly less relevant than the ones tailored for neuroticism (p<0.001) and agreeableness (p<0.01), and the feature-based explanation (p<0.001).

We recall here that the underlying recommended songs, on average, were similarly relevant, as they were based on a collaborative filtering algorithm and presented in a randomized order. As a result, we find that in our study, the design of the explanation may impact the *perceived fit* of a recommended item. Following [37], this may negatively impact the system's perceived usefulness and the users' satisfaction and trust.

5 Summary & Outlook

In this work, we explore the feasibility of personalizing explanations according to individual personality traits with the help of modern Generative AI technology. A statistical and linguistic analysis of personality-aware explanations suggests that tools like ChatGPT could be a reliable basis to create such explanations at scale. In the subsequent user study, we found that the different forms of explaining recommendations in a personality-aware form impact

[8]This is the set of participants who completed all necessary steps and passed an attention check during the study.

the users' perception of the explanations, and we also observed that not all forms of explaining were equally preferred.

In the music domain, the focus of our study, conscientiousness-oriented explanations might not be the best choice in general, even though they might work well in a particular listening context, such as podcasts. Notably, explanations aligned with low neuroticism were more frequently selected. This result may be closely tied to the emotional nature of music consumption, which is often driven by mood regulation rather than structured decision-making [19, 38].

Prior work has shown that people frequently turn to music to manage emotions [11]. This contrasts with conscientiousness, a trait linked to planning, structure, and task completion [45]. In our study, explanations reflecting this style emphasized order and productivity. While such framing may be appreciated in domains like education or productivity settings [47], music is usually emotional, not goal-oriented. In contrast, low-neuroticism explanations may offer users a sense of emotional stability, which aligns better with how people use music. These observations suggest that when designing explanations, it is important to consider not just the personality framing, but also the emotional and functional context in which the recommendation is being delivered.

Regarding study limitations, we recall that our population of study participants is biased towards females in a certain age range. Furthermore, since our pool of participants was all recruited from followers of a social book review platform, it has yet to be investigated if the findings regarding user perceptions would generalize to user groups with different profiles.

Overall, we see our work as a further step towards personality-aware explanations in recommender systems. As a part of our future research, our goal will be to study the connection between the personality traits of individual users and their preferences for certain personality-tailored explanations. One might, for example, assume that a highly open person might prefer explanations that emphasize terms that are related to openness in the explanation. Whether this is the true case and if it is thus possible to automatically select the "best" explanation type for each user, still has to be explored.

References

[1] Alan Agresti. 2002. *Categorical Data Analysis* (2nd ed.). John Wiley & Sons, Hoboken, NJ. doi:10.1002/0471249688

[2] Qurat Ul Ain, Mohamed Anime Chati, Mouadh Guesmi, and Shoeb Ahmed Joarder. 2022. A Multi-Dimensional Conceptualization Framework for Personalized Explanations in Recommender Systems. In *Joint Proceedings of the IUI 2022 Workshops co-located with the ACM International Conference on Intelligent User Interfaces (IUI 2022) (CEUR Workshop Proceedings, Vol. 3124)*. 11–23.

[3] Alaa Alslaity and Thomas Tran. 2020. The Effect of Personality Traits on Persuading Recommender System Users. In *Proceedings of the 7th Joint Workshop on Interfaces and Human Decision Making for Recommender Systems co-located with 14th ACM Conference on Recommender Systems (RecSys 2020) (CEUR Workshop Proceedings, Vol. 2682)*. 48–56.

[4] Liliana Ardissono, Anna Goy, Giovanna Petrone, Marino Segnan, and Pietro Torasso. 2003. Intrigue: Personalized Recommendation of Tourist Attractions for Desktop and Hand Held Devices. *Appl. Artif. Intell.* 17, 8-9 (2003), 687–714. doi:10.1080/713827254

[5] Krisztian Balog and Filip Radlinski. 2020. Measuring Recommendation Explanation Quality: The Conflicting Goals of Explanations. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20)*. 329–338. doi:10.1145/3397271.3401032

[6] Jotthi Bansal, Maya B Flannery, and Matthew H Woolhouse. 2021. Influence of personality on music-genre exclusivity. *Psychology of Music* 49, 5 (2021), 1356–1371. doi:10.1177/0305735620953611

[7] Shlomo Berkovsky, Ronnie Taib, and Dan Conway. 2017. How to Recommend? User Trust Factors in Movie Recommender Systems. In *Proceedings of the 22nd International Conference on Intelligent User Interfaces (IUI '17)*. 287–300. doi:10.1145/3025171.3025209

[8] Daniel Billsus and Michael J Pazzani. 1999. A personal news agent that talks, learns and explains. In *Proceedings of the 3rd Annual Conference on Autonomous Agents*. 268–275. doi:10.1145/301136.301208

[9] Mohamed Amine Chatti, Mouadh Guesmi, Laura Vorgerd, Thao Ngo, Shoeb Joarder, Qurat Ul Ain, and Arham Muslim. 2022. Is More Always Better? The Effects of Personal Characteristics and Level of Detail on the Perception of Explanations in a Recommender System. In *Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22)*. 254–264. doi:10.1145/3503252.3531304

[10] Robert B. Cialdini. 2004. The Science of Persuasion. *Scientific American Mind* 14, 1 (2004), 70–77. doi:10.1038/scientificamerican0201-76

[11] Terence Cook, Ashlin R. K. Roy, and Keith M. Welker. 2019. Music as an emotion regulation strategy: An examination of genres of music and their roles in emotion regulation. *Psychology of Music* 47, 1 (2019), 144–154. doi:10.1177/0305735617734627 arXiv:https://doi.org/10.1177/0305735617734627

[12] Sahraoui Dhelim, Nyothiri Aung, Mohammed Amine Bouras, Huansheng Ning, and Erik Cambria. 2022. A survey on personality-aware recommendation systems. *Art. Intell. Review* 55, 3 (2022), 2409–2454. doi:10.1007/s10462-021-10063-7

[13] Somayeh Fatahi, Mina Mousavifar, and Julita Vassileva. 2023. Investigating the effectiveness of persuasive justification messages in fair music recommender systems for users with different personality traits. In *Proceedings of the 31st ACM Conference on User Modeling, Adaptation and Personalization (UMAP '23)*. 66–77. doi:10.1145/3565472.3592958

[14] Bruce Ferwerda, Marko Tkalcic, and Markus Schedl. 2017. Personality Traits and Music Genres: What Do People Prefer to Listen To?. In *Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization (UMAP '17)*. 285–288. doi:10.1145/3079628.3079693

[15] Fatih Gedikli, Dietmar Jannach, and Mouzhi Ge. 2014. How should I explain? A comparison of different explanation types for recommender systems. *International Journal of Human-Computer Studies* 72, 4 (2014), 367–382. doi:10.1016/j.ijhcs.2013.12.007

[16] Sofia Gkika, Marianna Skiada, George Lekakos, and Panos E. Kourouthanassis. 2016. Investigating the Role of Personality Traits and Influence Strategies on the Persuasive Effect of Personalized Recommendations. In *Proceedings of the 4th Workshop on Emotions and Personality in Personalized Systems co-located with ACM RecSys 2016)*. 9–17.

[17] Lewis R. Goldberg. 1992. The Development Of Markers For The Big-Five Factor Structure. *Psychological Assessment* 4 (1992), 26–42. https://psycnet.apa.org/doi/10.1037/1040-3590.4.1.26

[18] Mouadh Guesmi, Mohamed Amine Chatti, Laura Vorgerd, Thao Ngo, Shoeb Joarder, Qurat Ul Ain, and Arham Muslim. 2022. Explaining User Models with Different Levels of Detail for Transparent Recommendation: A User Study. In *Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '22 Adjunct)*. 175–183. doi:10.1145/3511047.3537685

[19] Waldie E. Hanser, Tom F.M. ter Bogt, Annemieke J. M. Van den Tol, Ruth E. Mark, and Ad J. J. M. Vingerhoets. 2016. Consolation through music: A survey study. *Musicae Scientiae* 20, 1 (2016), 122–137. doi:10.1177/1029864915620264 arXiv:https://doi.org/10.1177/1029864915620264

[20] Mariam Hassanein, Wedad Hussein, Sherine Rady, and Tarek F. Gharib. 2018. Predicting Personality Traits from Social Media using Text Semantics. In *13th International Conference on Computer Engineering and Systems (ICCES)*. 184–189. doi:10.1109/ICCES.2018.8639408

[21] Jonathan L. Herlocker, Joseph A. Konstan, and John Riedl. 2000. Explaining Collaborative Filtering Recommendations. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work*. 241–250. doi:10.1145/358916.358995

[22] Oliver P. John and Sanjay Srivastava. 1999. The Big-Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. In *Handbook of Personality: Theory and Research*, Lawrence A. Pervin and Oliver P. John (Eds.). Vol. 2. Guilford Press, New York, 102–138.

[23] Rosemary Josekutty Thomas, Judith Masthoff, and Nir Oren. 2017. Adapting Healthy Eating Messages to Personality. In *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors*. 119–132. doi:10.1007/978-3-319-55134-0_10

[24] Bart P. Knijnenburg and Martijn C. Willemsen. 2015. Evaluating Recommender Systems with User Experiments. In *Recommender Systems Handbook*. Springer, 309–352. doi:10.1007/978-1-4899-7637-6_9

[25] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2020. Generating and Understanding Personalized Explanations in Hybrid Recommender Systems. *ACM Trans. Interact. Intell. Syst.* 10, 4 (2020), 1–40. doi:10.1145/3365843

[26] Peter J. Kwantes, Natalia Derbentseva, Quan Lam, Oshin Vartanian, and Harvey H.C. Marmurek. 2016. Assessing the Big Five personality traits with latent semantic analysis. *Personality and Individual Differences* 102 (2016), 229–233. doi:10.1016/j.paid.2016.07.010

[27] Lei Li, Yongfeng Zhang, and Li Chen. 2023. Personalized Prompt Learning for Explainable Recommendation. *ACM Trans. Inf. Syst.* 41, 4 (2023), 1–26. doi:10.1145/3580488

[28] Sandra C. Matz, Jacob D. Teeny, Srishti S. Vaid, Hannes Peters, Gabriel M. Harari, and Moran Cerf. 2024. The potential of generative AI for personalized persuasion at scale. *Scientific Reports* 14, 1 (2024), 4692. doi:10.1038/s41598-024-53755-0

[29] Alessandro B. Melchiorre and Markus Schedl. 2020. Personality Correlates of Music Audio Preferences for Modelling Music Listeners. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '20)*. 313–317. doi:10.1145/3340631.3394874

[30] Martijn Millecamp, Nyi Nyi Htun, Cristina Conati, and Katrien Verbert. 2019. To explain or not to explain: the effects of personal characteristics when explaining music recommendations. In *Proceedings of the 24th International Conference on Intelligent User Interfaces (IUI '19)*. 397–407. doi:10.1145/3301275.3302313

[31] Kimberly Sena Moore. 2013. A Systematic Review on the Neural Effects of Music on Emotion Regulation: Implications for Music Therapy Practice. *Journal of Music Therapy* 50, 3 (10 2013), 198–242. doi:10.1093/jmt/50.3.198

[32] Xia Ning and George Karypis. 2011. SLIM: Sparse linear methods for top-n recommender systems. In *Proceedings of the 11th IEEE International Conference on Data Mining (ICDM '11)*. 497–506. doi:10.1109/ICDM.2011.134

[33] Havva Alizadeh Noughabi, Behshid Behkamal, Fattane Zarrinkalam, and Mohsen Kahani. 2025. Persuasive explanations for path reasoning recommendations. *J. Intell. Inf. Syst.* 63, 2 (2025), 413–439. doi:10.1007/s10844-024-00896-3

[34] Ingrid Nunes and Dietmar Jannach. 2017. A Systematic Review and Taxonomy of Explanations in Decision Support and Recommender Systems. *User-Modeling and User-Adapted Interaction* 27, 3–5 (2017), 393–444. doi:10.1007/s11257-017-9195-0

[35] Michael P. Oakes and Malcolm Farrow. 2006. Use of the Chi-Squared Test to Examine Vocabulary Differences in English Language Corpora Representing Seven Different Countries. *Literary and Linguistic Computing* 22, 1 (2006), 85–99. doi:10.1093/llc/fql044

[36] Alexis Papadimitriou, Panagiotis Symeonidis, and Yannis Manolopoulos. 2012. A generalized taxonomy of explanations styles for traditional and social recommender systems. *Data Mining and Knowledge Discovery* 24, 3 (May 2012), 555–583. doi:10.1007/s10618-011-0215-0

[37] Pearl Pu, Li Chen, and Rong Hu. 2011. A User-centric Evaluation Framework for Recommender Systems. In *Proceedings of the 5th ACM Conference on Recommender Systems*. 157–164. doi:10.1145/2043932.2043962

[38] Suvi Saarikallio and Jaakko Erkkilä. 2007. The role of music in adolescents' mood regulation. *Psychology of Music* 35, 1 (2007), 88–109. doi:10.1177/0305735607068889

[39] Ítallo Silva, Leandro Marinho, Alan Said, and Martijn C. Willemsen. 2024. Leveraging ChatGPT for Automated Human-centered Explanations in Recommender Systems. In *Proceedings of the 29th International Conference on Intelligent User Interfaces (IUI '24)*. 597–608. doi:10.1145/3640543.3645171

[40] Xiangguo Sun, Bo Liu, Jiuxin Cao, Junzhou Luo, and Xiaojun Shen. 2018. Who Am I? Personality Detection Based on Deep Learning for Texts. In *2018 IEEE International Conference on Communications (ICC)*. 1–6. doi:10.1109/ICC.2018.8422105

[41] Nava Tintarev and Judith Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction* 22, 4–5 (2012), 399–439. doi:10.1007/s11257-011-9117-5

[42] Nava Tintarev and Judith Masthoff. 2022. Beyond Explaining Single Item Recommendations. In *Recommender Systems Handbook*, Francesco Ricci, Lior Rokach, and Bracha Shapira (Eds.). Springer US, New York, NY, 711–756. doi:10.1007/978-1-0716-2197-4_19

[43] Paul D. Trapnell and Jerry S. Wiggins. 1990. Extension of the Interpersonal Adjective Scales to include the Big Five dimensions of personality. *Journal of Personality and Social Psychology* 59, 4 (1990), 781–790. doi:10.1037/0022-3514.59.4.781

[44] Sahil Verma, Chirag Shah, John P. Dickerson, Anurag Beniwal, Narayanan Sadagopan, and Arjun Seshadri. 2023. RecXplainer: Amortized Attribute-based Personalized Explanations for Recommender Systems. arXiv:2211.14935 [cs.IR] https://arxiv.org/abs/2211.14935

[45] Yvonne Vermetten, Hans Lodewijks, and Jan Vermunt. 2001. The Role of Personality Traits and Goal Orientations in Strategy Use. *Contemporary Educational Psychology* 26 (05 2001), 149–170. doi:10.1006/ceps.1999.1042

[46] Yongfeng Zhang and Xu Chen. 2020. Explainable Recommendation: A Survey and New Perspectives. *Foundations and Trends® in Information Retrieval* 14, 1 (2020), 1–101. doi:10.1561/1500000066

[47] Åge Diseth. 2003. Personality and approaches to learning as predictors of academic achievement. *European Journal of Personality* 17, 2 (2003), 143–155. doi:10.1002/per.469 arXiv:https://doi.org/10.1002/per.469