Effects of Representation Nudges on the Perception of Playlist Recommendations

Shah Noor Khan[1[0009-0006-4899-2987]], Eelco Herder[1[0000-0003-0107-299X]], and Diba Kaya[2[]]

[1] [1]Utrecht University, Utrecht, The Netherlands
{s.n.khan1, e.herder}@uu.nl
[2] [2]Jstor, United States
dibakaya@gmail.com

Abstract. Music plays a significant role in social interactions across cultures, carrying diverse meanings and serving various personal and global purposes. Therefore, it is crucial that music platforms ensure that users from diverse groups feel recognized, fairly treated, and satisfied with their recommendations. This study examines how "representation" nudges influence playlist selection and user attitudes towards different minority groups. The results indicate that users appreciated the labels, provided that they recognized the associated user group as a minority that needs attention. In some cases, users changed their playlists accordingly, particularly when they are in line with their personal preferences.

Keywords: Music recommendation · Playlists · Minority representation · Nudges

1 Introduction

Music is an integral part of social interactions and experiences in diverse cultures [36]. It carries significant social meanings and serves various functions, from fostering self-awareness and creating social bonds to conveying emotions [18]. Music is a potent medium for expressing local concerns, with teenagers especially using their music preferences to shape and define their identities [30, 25].

Recommender Systems (RS) process large volumes of choices to suggest the most appropriate and relevant items to users. However, these systems tend to favor certain items, recommending them more frequently, while other items receive less attention [22]. As a result, the group with the least coherence to popular tastes is most affected, often receiving poorer recommendations [1].

Research shows that users who prefer non-mainstream music have more extensive profiles and listen to a wider variety of artists than those who favor mainstream music. This highlights the need to improve recommendations for these users [22]. Recommending less popular, long-tail items is also beneficial to diversify recommendations for mainstream users and to increase exposure of lesser known artists and songs [3, 29]. Platforms like Music recommender system

(MRS) use techniques such as nudging [11] to subtly guide user decisions toward better outcomes without enforcing specific choices [35].

For this study, we tailored playlists for user dimensions like age, origins, gender, and sexual orientation, and used representation nudges to inform users – who may or may not belong to the associated user groups – about these dimensions, in order to assess their acceptance. These dimensions have been previously explored in evaluating biases in recommender systems [12, 14]. Also, the study aimed to determine whether the nudges associated with various minority groups influenced the participants' playlist preferences.

2 Literature Review

In RS, fairness and equity are increasingly considered important. Research [21, 32] on fairness often defines fairness uniformly across all groups. However, it is essential to recognize that fairness should be defined differently for different groups [6]. Freeman et al. [16] view fairness as a dynamic state, where current relations between advantaged and disadvantaged groups are not considered as a given, but as dynamics that should be explicitly shaped by system decisions. Researchers have recently explored rich sub-group fairness, aiming for fairness across multiple intersecting subgroup categories [21].

Within recommender systems, nudges are interface elements that aim to steer user choices in directions that are considered more desirable [20]. Nudging has proven effective in various areas, including sustainable transportation [2], energy conservation [33], green transportation [5], eco-friendly clothing [8], healthier eating [34, 13], and privacy awareness [28]. In music recommendations, for instance, nudges may help users explore new genres, encouraging them to go beyond their usual preferences [23].

Playlists are vital to online music listening, with most users creating them for personal and shared use. Platforms like Spotify offer tools for creating and curating playlists based on user preferences. The curation of playlists highlighting specific themes serves as a form of self-expression, a feature of identity, and contributes to the expansion of public culture [26, 27]. This curation can also raise awareness about inequality and help unite minority communities [37]. Another study demonstrates that identity plays a crucial role in the creation of playlists on streaming platforms [17].

Studies [15, 24] highlight concerns about gender balance and bias, with women often receiving less accurate recommendations. Also, recommending ethnic or region-specific music is challenging due to its poor fit with Western-based methods and small user bases [7, 9]. Research [10] examines how LGBTQ-themed playlists influence and reflect LGBTQ identities and cultures.

In our current study, we aim to apply nudging strategies to highlight minority groups in recommender systems, thereby fostering a more inclusive platform. This research could potentially alter user directions and increase awareness of minority groups within the majority.

3 Methodology

This section describes the method and study design. We utilized the vignettes method [4], which is effective in HCI studies. This technique involves presenting participants with specific scenarios or stories and then asking them questions to understand their reactions, attitudes, or decision-making processes. Participants were provided with different scenarios introducing nudges and were asked to select playlists and answer questions related to their choices in the study.

The purpose of the study was to explore whether playlist adaptations affect participants' choices in terms of playlist selection. There were two rounds in this study. In each round, participants were presented with four sets of three playlists, with each playlist consisting of 15 songs. Each set of three playlists was presented on a separate screen, and in each set, there was one playlist with a specific equity-focused optimization and two generic, base-line playlists.

All playlists included a balanced mix of songs selected from the current charts[3], popular songs from the 2010s[4], and popular R&B songs[5]. Additionally, each optimized playlist contained 4 songs selected for users from a particular age group, origin, gender, or orientation, selected from reputable national and international song collections: a 40+ Retro Playlist with popular classic rock songs[6], a Europeanized Playlist including Dutch, German and French songs[7], a Gender-Balanced Playlist featuring Power Women songs[8], and a Rainbow Playlist boasting songs popular among the queer community[9].

We created an interactive website for this study[10], with the aim to entertain the participants and to solicit their reactions. The study was carried out either in person or online via Teams, with the experimenter guiding the participants through the study. Participants consented to be recorded.

In the first round, participants were consecutively presented with four web pages, each containing a set of three playlists. Their task was to simply select the playlist they preferred from the set of three. An example selection screen is displayed in Figure 1.

Participants were not given any indication about possible playlist interventions and were asked to think aloud during the process. Each interface contained one optimized playlist and in order to keep conversations comparable, the order of optimized playlists and associated user groups (age, origin, gender,

[3] https://www.top40.nl/top40/2024/week-4
[4] https://open.spotify.com/playlist/6JpGyIjFccZ6COElPnDZKu
[5] https://thefortyfive.com/opinion/best-r-and-b-songs-of-all-time/
[6] https://www.arrow.nl/wp-content/uploads/2023/10/Arrow-Classic-Rock-500-2023-100-11.pdf
[7] https://www.nporadio2.nl/nieuws/npo-radio-2/31f4a2a2-32fc-4e75-88f1-b6e73a283f5d/dit-is-de-koninklijke-500-van-2022, https://www.muziekweb.nl/en/Link/HCX1490/De-40-bekendste-Duitse-liedjes, https://www.top40.nl/overig-nieuws-de-grootste-franstalige-hit
[8] https://www.seventeen.com/celebrity/music/g22878449/girl-power-songs/
[9] https://www.timeout.com/music/best-gay-songs
[10] https://www.projects.science.uu.nl/ics-playlistselection/

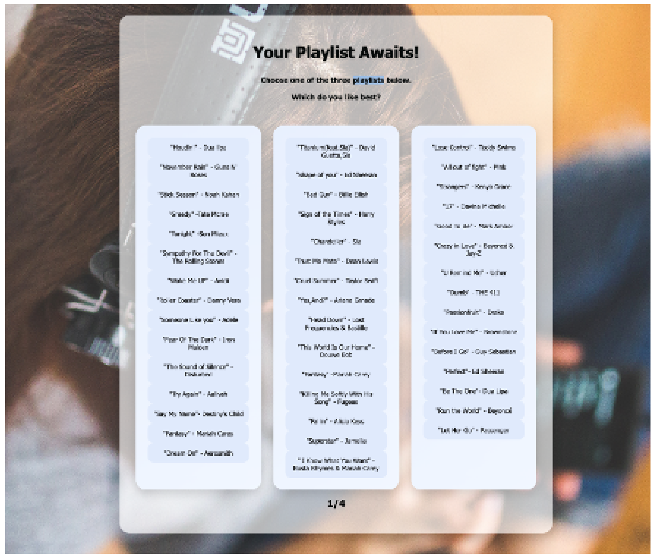

Fig. 1. Playlist selection in the first round. All playlists, including the optimized one, are presented equally, in randomized order.

orientation) was fixed, ordered by the expected increasing controversy. Within each interface, the order of the playlists was randomized to avoid middle-option bias [31], but the altered playlist remained the same.

In the second round, four types of labels were used, each referring to the user group dimensions mentioned above. The labels were intended to serve as nudges to raise attention for the suggested disadvantages that user groups may experience within the platform. The labels of the optimized playlists had a different pastel background color compared to the other two playlists. The pastel colors were used to reduce any bias caused by color. In addition, the songs associated with the user group dimensions were highlighted with a different pastel background color, demonstrating how these user groups could be better accommodated. In Figure 2, there is an example of the optimized playlist, the label suggests that it is a Europeanised playlist and the highlighted songs with different background are the songs associated with this dimension.

At the end of the second round, participants were asked questions related to how comfortable they are with the optimizations, how well they rated the adaptations to the user groups, as well as their familiarity and association with each of the four user groups. Finally, participants were asked to comment upon the optimized playlist and they were invited to change their choice or to indicate that they would stick to their initial choice with this additional knowledge.

Fig. 2. An optimized playlist in the second round, with user group label and highlighted songs associated with this user group.

4 Results

The study was qualitative in nature and, since it was a preliminary study, we initially aimed to have 10 participants. However, after 8 participants, we found that we had reached the intended saturation level.

4.1 User Interaction and Playlist Choices

The playlist selection in the first round and the playlist confirmation or change in the second round served as a basis for the interviews, reported in the next subsection. Therefore, we start with a summary of these results. The numbers presented are indicative and used for interpretation. The participants saw the playlists in random order, but for convenience, in this section, the optimized playlist is always Playlist 1.

Age Dimension (40+ Retro): In both rounds, Playlist 1 and Playlist 2 were picked four times each, showing no particular preference. Playlist 3 was not selected at all. Apparently, the optimization did not affect participants' choices.

Origin Dimension (Europeanized): In the first round, Playlist 1 was chosen twice, and Playlist 3 was chosen six times. After having seen the nudges in round 2, Playlist 1 was chosen six times, and Playlist 3 was chosen twice. This indicates that awareness of the optimization influenced participants to choose Playlist 1 more.

Gender Dimension (Power Women): In the first round, Playlist 1 was chosen five times, and Playlist 3 was chosen three times. After having seen the representation nudges, all participants had chosen Playlist 1, apparently indicating a strong association with the gender dimension.

Orientation Dimension (Rainbow): In the first round, Playlist 1 was chosen once, and Playlist 3 was chosen seven times. After the nudges have been shown, Playlist 1 was chosen five times, and Playlist 3 was chosen three times: five participants changed their choices due to awareness of the optimization.

Overall, the results indicate that the optimized playlists influenced participants' choices, especially for the gender, European, and sexual orientation dimensions. The age dimension optimization had no apparent effect.

4.2 Interviews

All interviews were transcribed and coded. This section is structured based on the most important and frequent codes.

Initial selection criteria: Participants' playlist selections were influenced by factors that included familiarity with artists and songs, personal preferences, situational needs, and recommendations. For instance, P1 and P4 prefer familiar artists like Kensington, Ed Sheeran and Dua Lipa, while P1 and P3 avoid Justin Bieber. Further, P2 and P4 chose playlists based on specific goals or situations, whereas P1 and P7 rely on Spotify's recommendations. Nostalgia also plays a role, with P1 and P4 mentioning songs from their teenage years.

Affiliation with user groups/dimensions: Participants showed reluctance to choose the optimized playlist for the 40+ demographic, as they felt this playlist did not resonate with their personal identity or music tastes. For example, P3 and P8 noted that such playlists felt irrelevant to them, with P7 humorously remarking that a 40+ label made them feel old, despite being only 30.

There was a strong tendency towards supporting European artists, with participants highlighting the importance of promoting local talent over dominant American artists. P6 and P4 mentioned choosing playlists featuring European artists to align with their tastes and support regional music scenes.

The discussion also highlighted benefits of playlists optimized for queer or female artists, seeing these as opportunities to explore diverse perspectives and support underrepresented groups. P7, who does not identify as queer, still expressed a desire to support queer artists, because they enjoy the positive and self-affirming messages in their music.

In conclusion, while demographic labels on playlists can influence music selection, personal identity and significance related to those identities play a significant role in playlist selection. Participants preferred playlists that align with their cultural and regional affiliations and that promote diversity and inclusion, but at the same time showed hesitation in selecting playlists for a age group with whom they do not identify and which don't consider as relevant.

Effects of the optimization: Participants described varied responses to the effects of playlist optimizations, reflecting their individual preferences and decision-making processes. P1 and P6 stated that awareness of the optimizations did not significantly change their choice. They mentioned that they would have chosen the playlist quicker with the label that served as representation nudge, but would still go for the same one regardless of the label.

P2, P5, and P7 found the labels interesting, making the playlists seem less generic. They liked the categorization and felt it helped them discover new music. P5 felt that the labels provided an extra reason to listen to the playlist and made them more aware of inclusivity. P7 appreciated the variety and the opportunity to discover new music from female artists, mentioning that the labels influenced their initial choice, but would still consider the songs in the playlist.

P3 and P4 had mixed feelings about the labels. P3 shared that while they usually choose based on the title of the playlist, the additional information made them consider the gender balance and variety within the playlist. They found the new dimension of variety appealing and felt it increased the likelihood of choosing the playlist. P4 felt that while the labels provided a new perspective, they did not always resonate with their personal preferences. P8 felt that the labels did not resonate with them personally and preferred playlists that focused on the music genre rather than the audience's age or underrepresented groups.

Acceptance of highlighting minorities: Several participants expressed a desire for the inclusion of various categories to cater to different tastes and groups. P1 stated that being part of a minority group influences their interest in such playlists, while P6 emphasized the potential for these playlists to become mere token gestures if not implemented sincerely. The idea of supporting one's minority group through algorithmic boosts was seen positively by P5.

There was also a concern about potential backlash from highlighting minority groups, especially from right-wing movements. Despite this, P2 felt that representation is crucial and that having visible categories for minority communities is beneficial. P3 was interested in seeing if exposure to minority-focused playlists could change perceptions among those not identifying with those communities.

5 Discussion

In the first round of playlist selection without any labels, the majority of participants mainly based their choices on familiarity with artists and their liking towards the songs. The context in which they were listening also influenced their selections. These results are expected and in line with current literature [19]. In the second round, the participants were shown the optimization labels, along with the highlighting of the optimized tracks. This affected their selection criteria, as described in the interview results.

Our participants, all between 20 and 30, found it hard to associate themselves with the optimized playlist for 40+ users, also because they did not consider this group a user group that lacks representation. By contrast, the European dimension did prompt participants to change their choices, as they are aware of the issues faced by these artists and consider them a minority on the platform.

For the dimension of gender and orientation, there was a significant effect as well. Since the participants were already familiar with these groups being discriminated, they felt that promoting these groups would create a more positive and inclusive atmosphere. Their familiarity with these groups is also a result of these groups being discussed on a broader level in the current society.

However, the participants' decision to select a different playlist was affected by personal preferences as well. Although a couple of participants considered the minority representation an important aspect, they still preferred to go with the option they originally selected. This may be interpreted as a potential *conflict of interest* between the wish (or social desirability) to appreciate values like inclusivity on the one hand and the wish to enjoy songs of their own preference on the other hand – in cases where these two factors do not align.

Furthermore, participants emphasized that superficial labeling would not be helpful. There was a mention of backlash as well from groups that do not like minorities being represented, but still considered visibility to be important and beneficial, as this could also help in changing perceptions among such users.

6 Conclusion

In this study, we aimed to investigate whether optimized playlists and associated labels for four different minority groups (age group, origin, gender and orientation) had any effect on the choice selection of participants. The results indicated that these labels did influence participants' choices, but only if they identified with or sympathized with the respective minority group, emphasizing the importance of meaningful labels over superficial ones. The second aspect of the study focused on whether participants would be comfortable with a platform that highlights minority preferences. Participants indeed showed positive attitudes, although they raised questions about potential backlash.

This preliminary study, being qualitative in nature, could be expanded into a more quantitative study by using the interface as a survey tool to gather responses from a larger sample group, in order to analyze patterns in playlist selection. Systematic variations and planned comparisons could also provide specific strategies to solve potential conflicts of interest between users wanting to show their sympathy with specific non-mainstream user groups – by accepting (or even embracing) the associated songs and genres in their choice architecture – and the desire to still have playlists and recommendations optimized for their own preferences.

References

1. Abdollahpouri, H., Mansoury, M., Burke, R., Mobasher, B.: The connection between popularity bias, calibration, and fairness in recommendation. In: Proceedings of the 14th ACM conference on recommender systems. pp. 726–731 (2020)
2. Anagnostopoulou, E., Urbančič, J., Bothos, E., Magoutas, B., Bradesko, L., Schrammel, J., Mentzas, G.: From mobility patterns to behavioural change: leveraging travel behaviour and personality profiles to nudge for sustainable transportation. Journal of Intelligent Information Systems **54**, 157–178 (2020)
3. Anderson, C.: Debating the long tail. Harvard Business Review (2008)
4. Atzmüller, C., Steiner, P.M.: Experimental vignette studies in survey research. Methodology (2010)
5. Bothos, E., Apostolou, D., Mentzas, G.: Recommender systems for nudging commuters towards eco-friendly decisions. Intelligent Decision Technologies **9**(3), 295–306 (2015)
6. Burke, R., Mattei, N., Grozin, V., Voida, A., Sonboli, N.: Multi-agent social choice for dynamic fairness-aware recommendation. In: Adjunct Proceedings of the 30th ACM Conference on User Modeling, Adaptation and Personalization. pp. 234–244 (2022)
7. Cornelis, O., Lesaffre, M., Moelants, D., Leman, M.: Access to ethnic music: Advances and perspectives in content-based music information retrieval. Signal Processing **90**(4), 1008–1031 (2010)
8. Cossatin, A.G., Mauro, N., Ardissono, L.: Promoting green fashion consumption through digital nudges in recommender systems. IEEE Access (2024)
9. Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U., Gupta, S., He, Y., Lambert, M., Livingston, B., et al.: The youtube video recommendation system. In: Proceedings of the fourth ACM conference on Recommender systems. pp. 293–296 (2010)
10. Dhaenens, F., Burgess, J.: 'press play for pride': The cultural logics of lgbtq-themed playlists on spotify. new media & society **21**(6), 1192–1211 (2019)
11. Dolgopolova, I., Li, B., Roosen, J.: Do nudges matter? consumer perception and acceptance of recommender systems with different types of nudges. In: 3rd FAccTRec Workshop: Responsible Recommendation at The ACM Conference Series on Recommender Systems RECSYS 2020 (2020)
12. Ekstrand, M.D., Tian, M., Azpiazu, I.M., Ekstrand, J.D., Anuyah, O., McNeill, D., Pera, M.S.: All the cool kids, how do they fit in?: Popularity and demographic biases in recommender evaluation and effectiveness. In: Conference on fairness, accountability and transparency. pp. 172–186. PMLR (2018)
13. El Majjodi, A., Starke, A.D., Trattner, C.: Nudging towards health? examining the merits of nutrition labels and personalization in a recipe recommender system. In: Proceedings of the 30th ACM conference on user modeling, adaptation and personalization. pp. 48–56 (2022)
14. Farnadi, G., Kouki, P., Thompson, S.K., Srinivasan, S., Getoor, L.: A fairness-aware hybrid recommender system. arXiv preprint arXiv:1809.09030 (2018)
15. Ferraro, A., Serra, X., Bauer, C.: Break the loop: Gender imbalance in music recommenders. In: Proceedings of the 2021 conference on human information interaction and retrieval. pp. 249–254 (2021)
16. Freeman, R., Zahedi, S.M., Conitzer, V.: Fair social choice in dynamic settings. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI). pp. 4580–4587. International Joint Conferences on Artificial Intelligence Marina del Rey, CA (2017)

17. Hagen, A.N.: The playlist experience: Personal playlists in music streaming services. Popular Music and Society **38**(5), 625–645 (2015)
18. Hallam, S., Cross, I., Thaut, M.: Oxford handbook of music psychology. Oxford University Press (2009)
19. Jannach, D., Lerche, L., Jugovac, M.: Item familiarity effects in user-centric evaluations of recommender systems. In: RecSys Posters (2015)
20. Jesse, M., Jannach, D.: Digital nudging with recommender systems: Survey and future directions. Computers in Human Behavior Reports **3**, 100052 (2021)
21. Kearns, M., Neel, S., Roth, A., Wu, Z.S.: Preventing fairness gerrymandering: Auditing and learning for subgroup fairness. In: International conference on machine learning. pp. 2564–2572. PMLR (2018)
22. Kowald, D., Schedl, M., Lex, E.: The unfairness of popularity bias in music recommendation: A reproducibility study. In: Advances in Information Retrieval: 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14–17, 2020, Proceedings, Part II 42. pp. 35–42. Springer (2020)
23. Liang, Y., Willemsen, M.C.: Promoting music exploration through personalized nudging in a genre exploration recommender. International Journal of Human–Computer Interaction **39**(7), 1495–1518 (2023)
24. Mansoury, M., Abdollahpouri, H., Smith, J., Dehpanah, A., Pechenizkiy, M., Mobasher, B.: Investigating potential factors associated with gender discrimination in collaborative recommender systems. In: The thirty-third international flairs conference (2020)
25. North, A.C., Hargreaves, D.J.: Music and adolescent identity. Music education research **1**(1), 75–92 (1999)
26. Potter, J.: Digital media and learner identity: The new curatorship. Springer (2012)
27. Potter, J., Gilje, Ø.: Curation as a new literacy practice (2015)
28. Salem, R.B., Aïmeur, E., Hage, H.: A nudge-based recommender system towards responsible online socializing. In: OHARS@ RecSys. pp. 23–39 (2020)
29. Shani, G., Gunawardana, A.: Evaluating recommendation systems. Recommender systems handbook pp. 257–297 (2011)
30. Simon, F.: Sound effects: Youth, leisure and the politics of rock 'n'roll. Constable, London (1981)
31. Simons, A., Weinmann, M., Tietz, M., vom Brocke, J.: Which reward should i choose? preliminary evidence for the middle-option bias in reward-based crowdfunding (2017)
32. Sonboli, N., Eskandanian, F., Burke, R., Liu, W., Mobasher, B.: Opportunistic multi-aspect fairness through personalized re-ranking. In: Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization. pp. 239–247 (2020)
33. Starke, A., Willemsen, M., Snijders, C.: Effective user interface designs to increase energy-efficient behavior in a rasch-based energy recommender system. In: Proceedings of the eleventh ACM conference on recommender systems. pp. 65–73 (2017)
34. Starke, A.D., Asotic, E., Trattner, C., Van Loo, E.J.: Examining the user evaluation of multi-list recommender interfaces in the context of healthy recipe choices. ACM Transactions on Recommender Systems **1**(4), 1–31 (2023)
35. Thaler, R.H., Sunstein, C.R., Balz, J.P.: Choice architecture. The behavioral foundations of public policy (2014)
36. Turino, T.: Music as social life: The politics of participation. University of Chicago Press (2008)

37. Wargo, J.M.: # donttagyourhate: Reading collecting and curating as genres of participation in lgbt youth activism on tumblr. Digital Culture & Education **9**(1) (2017)