

The Importance of Cognitive Biases in the Recommendation Ecosystem: Evidence of Feature-Positive Effect, Ikea Effect, and Cultural Homophily

Markus Schedl^{1,2,*}, Oleg Lesota¹ and Shahed Masoudian¹

¹Institute of Computational Perception, Johannes Kepler University Linz (JKU), Altenberger Straße 69, A-4040 Linz, Austria

²Human-centered AI Group, AI Lab, Linz Institute of Technology (LIT), Altenberger Straße 69, A-4040 Linz, Austria

Abstract

Cognitive biases have been studied in psychology, sociology, and behavioral economics for decades. Traditionally, they have been considered a negative human trait that leads to inferior decision making, reinforcement of stereotypes, or can be exploited to manipulate consumers, respectively. Lately, there has been growing interest in AI research to better understand the influence of such biases in classification, search, and also recommendation tasks. We argue that cognitive biases manifest in different parts of the recommendation ecosystem and in various components of the recommendation pipeline, including input data (such as ratings or side information), recommendation algorithm or model (and consequently recommended items), and user interactions with the system. More importantly, we contest the traditional detrimental perspective on cognitive biases and claim that certain cognitive biases can be beneficial when accounted for by recommender systems. Concretely, we provide empirical evidence that feature-positive effect, Ikea effect, and cultural homophily can be observed in the context of recommender systems, and discuss their potential for exploitation. In three small experiments covering recruitment and entertainment domains, we study the pervasiveness of the aforementioned biases. We ultimately advocate for a prejudice-free consideration of cognitive biases to improve user and item models as well as recommendation algorithms.

Keywords

Psychology, Cognition, Feature-Positive Effect, Ikea Effect, Cultural Homophily, Declinism, Halo Effect, Primacy Effect, Recency Effect, Position Bias, Empirical Studies, Simulation Study

1. Introduction and Background

“Music used to be better in the 1980s when I was young.”

“The cookies I baked are much tastier than the ones I bought.”

“I only remember the last items on my to-do list.”

“He is such a great actor; I am sure those nasty accusations against him are made up.”

These are examples of common cognitive biases, respectively, declinism, Ikea effect, recency effect, and halo effect. Such biases have been studied in psychology and sociology for decades. In psychology, they are commonly defined as systematic perceptual deviations of the individual from rationality and objectivity, cognition, judgment, or decision making, which often happens unconsciously [1, 2]. In sociology, they typically refer to collective prejudices of a society that favor one group’s values, norms, and traditions over others [3, 4].

Historically, cognitive biases have been regarded as a negative characteristic of humans, which lead to inferior decision making, reinforce stereotypes, and may even cause severe systematic errors and harm [5]. More recently, psychologists have started to acknowledge the positive effects of certain cognitive biases, e.g., to improve the efficiency of human learning and decision making [6, 1]. In the field of machine learning, the study of cognitive biases has played a minor role so far. Only lately, some ideas to leverage cognitive biases for model training, e.g., to improve their generalization capabilities or foster

IntrS’24: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, October 18, 2024, Bari (Italy)

*Corresponding author.

✉ markus.schedl@jku.at (M. Schedl); oleg.lesota@jku.at (O. Lesota); shahed.masoudian@jku.at (S. Masoudian)

🌐 <https://hcai.at> (M. Schedl)

🆔 0000-0003-1706-3406 (M. Schedl)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

ethical machine behavior emerged [7]. Narrowing down the scope to search and ranking, in information retrieval, some preliminary research on cognitive biases has been conducted recently [8, 9, 10]. Existing research, however, has focused on detecting some cognitive biases and assessing their influence on search behavior rather than leveraging them to improve retrieval algorithms.

In recommender systems (RSs) research, some psychologically grounded human biases have been studied in the past, e.g., primacy and recency effects in peer recommendation [11] as well as risk aversion and decision biases in product recommendation [12]. However, despite their importance, cognitive biases in the context of recommendation have received surprisingly little attention over the past few years. Nor are we aware of any systematic investigation of their manifestations at different stages of the recommendation process. This is particularly astonishing because research in RSs has historically been inspired by psychological theories, models, and empirical evidence on human decision making [13].

To narrow this research gap, we advocate for a *holistic examination of cognitive biases within the recommendation ecosystem*. And we take first steps in this direction in the paper at hand. Overall, we argue for studying their potential manifestations in different *components* of the system, at different *stages* of the recommendation process, and from the perspective of different *stakeholders*. Furthermore, we aim to evaluate and harness the positive effects these biases may have, with the goal of enhancing user and item models, as well as refining recommendation algorithms.

In the following, we briefly introduce a selection of cognitive biases that we believe deserve a more thorough investigation in the context of RSs (Section 2). We then present empirical evidence of some of these biases and showcase how they may influence recommendation (Section 3). Ultimately, we argue for a much-needed (re-)consideration of cognitive biases and point to important directions this could take (Section 4).

2. Cognitive Biases in Recommendation

Extensive research in psychology, sociology, and economics has revealed a plethora of cognitive biases [1, 2]. They relate to how humans perceive, process, store, and retrieve information, involving the cognitive and neurological processes of individuals and even whole societies. While not all cognitive biases directly apply to RSs, many of them influence user behavior and decision-making processes. As a result, these biases affect users' interactions with items (e.g., ratings or consumption patterns) and with RSs in general (e.g., use of different functionalities provided by the system's interface). However, only a few cognitive biases have been studied in the context of RSs. In the following, we introduce some, point to related work, and provide a working definition in the context of the recommendation ecosystem.¹

Feature-Positive Effect [15, 16]: Humans better realize and put more emphasis on things that are present than on things that are absent. In the context of RSs, we argue that this effect plays a crucial role in explainability and fairness [17]. For instance, through counterfactuals, one could show users which (maybe better-suited) items would have been recommended to them if they had different traits. We demonstrate this effect in Section 3.1.

Ikea Effect [18, 19]: The more effort a person invested in something, the more they will value it. This effect owes its name to a study that found participants were willing to pay a much higher price for furniture they had helped build than for ready-made items [18]. It can be thought of as the desire of humans to justify their efforts. In a RS context, we assume the existence of similar effects when users create their own item collections (e.g., songs or videos in a playlist, visited places) in contrast to interacting with those created by others or by a recommendation algorithm. We provide evidence for this effect in Section 3.2.

Homophily [20, 21]: Homophily refers to the fact that individuals tend to form connections with others who have similar characteristics (e.g., age, culture, or religion) more often than with people having different traits. While it is a well-studied phenomenon in sociology [21] and even evolutionary

¹This is, by no means, meant to be an exhaustive list, but a biased selection of the authors. For a more in-depth discussion about cognitive biases, we refer to [14, 13].

biology [22], in RSs, homophily has not been extensively studied from a cognitive psychological perspective.² Evidence of cultural homophily in the music domain has been found in [23], where music listeners of certain countries (e.g., Brazil, Sweden, and Germany) showed a preference to listen to domestic music artists. Vice versa, Brazilian, Polish, and Russian artists were found to be predominantly liked by their domestic audience. In Section 3.3, we provide additional evidence that such effects can be observed and influenced by the recommendation algorithm.

Declinism [24, 25]: The perception that the world or society is declining, i.e., things get worse over time. This has been shown to be (partly) the result of *rosy retrospection* – humans’ tendency to remember the past as more positive as it actually was [26]. Declinism bias can be studied in the context of RSs, by considering side information on items, e.g., emotions reflected in song lyrics have become more negative over the last decades [27]. We hypothesize that such trends can also be observed in longer-term historic interaction data, i.e., users tend to interact more frequently over time with items that are negatively emotion-laden. Identifying and formalizing such trends could also be used to adjust recommendations to counteract (or amplify) this bias.

Contrast Effect [28]: If two items are shown in the vicinity in the recommendation list of a user – one with exceptional, the other with medium utility – the latter will appear much less appealing to the user, even if it is still a reasonable choice.

Anchoring [29]: As a variant of the contrast effect, anchoring refers to the fact that humans often overemphasize the piece of information they are exposed to first. RS providers can exploit this effect to offer a hook to their users, e.g., showing them a highly-priced item first, making them (unconsciously) believe that subsequent items are cheap even if they are still overpriced [30]. This use is also referred to as *decoy effect* [31].

Conformity Bias [32]: Providing users with evidence of previous interactions or ratings by others impacts their own consequent behavior. For instance, showing users (artificial) ratings before asking them to provide their own results in their ratings being closer to the initially presented ones [32]. Users were also found more likely to click on an item if they saw that many others have done so [33]. Furthermore, recent research reported in [34] has studied two types of conformity bias in RSs: informational conformity (belief that one’s peer group has superior knowledge) and normative conformity (desire for social approval). The authors also propose a model to disentangle individual users’ self-interest and conformity behavior.

Primacy and Recency Effect (Position Bias) [35, 11]: The position at which items occur in a recommendation list influences the probability that users will interact with them. Users are more likely to interact with items appearing at the beginning (*primacy effect*) and at the end (*recency effect*) of a ranking or sequence of recommendations. In some specific domains, *position bias* (or *serial position effects*) have been observed already, e.g., users were more likely to vote for recommended science stories shown to them first and last [11]. Notably, the primacy effect was much stronger than the recency effect in this case. In product recommendation, an influence of item position on users’ ability to remember product characteristics and users’ inclination to select certain products were observed [36]. Recency effects have also been formalized in the cognitive architecture ACT-R [37], which has been used to predict repeated consumption behavior [38] and to increase diversity and explainability of RSs [39].

Halo Effect [40, 41]: One’s overall impression or specific perceived traits of a person influence the perception of other characteristics of that person. For instance, it has been shown that physically attractive people tend to be perceived as more intelligent and having more positive personality traits than less attractive ones [42]. This overly positive perception overshadows the negative traits of the person. In the context of RSs, we assume that similar effects occur, e.g., in slate-, playlist-, or basket recommendation, where the recommended collection of items as a whole may be perceived as more or less favorable depending on a single trait of one salient item (or its creator) in the recommended item collection. If evidence for this can be found, RS providers could, for instance, implement a mechanism to push content created by underrepresented producers by adding it to a collection with items whose

²This is even more surprising because collaborative filtering algorithms leverage this concept, by exploiting similarity in preferences or interaction behaviors.

halo will extend to those injected items, thereby improving the latter’s exposure and overall fairness.

It is important to note that cognitive biases usually do not occur in isolation. Instead, in the recommendation process, they are often intertwined and related through feedback loops. For instance, position bias influences the probability of users clicking on a recommended item, which affects the interaction data used to train or update the recommendation model. Also, primacy-, contrast-, and anchoring effects are often at work jointly.

3. Empirical Findings

To showcase manifestations of cognitive biases in the context of RSs, we present several preliminary investigations. Among the many existing cognitive biases, we select feature-positive effect (Section 3.1), Ikea effect (Section 3.2), and cultural homophily (Section 3.3) because they received only little attention so far by the RS community. Furthermore, this selection of biases and choice of experimental design covers different recommendation components and stages, respectively, (content-based) training data, user interactions, and recommendation outcomes. Our experiments cover the domains of recruitment (candidate recommendation) and entertainment (music recommendation).

3.1. Feature-Positive Effect

The feature positive effect (FPE) is known as the tendency of learning organisms to better detect the presence of stimuli (e.g., *if p then q*) rather than their absence (e.g., *if p then not q* or *if not p then q*) [16, 43]. This concept can be relevant in job and candidate RSs, as these systems also focus on what is present in a job ad to determine an applicant’s relevance, potentially overlooking what is missing in the job ad.

To investigate the presence of the FPE in candidate RSs (for given job ads), we developed a content-based RS using the Distil-RoBERTa cross-encoder model, similar to Kumar et al.’s work on detecting biases in job RSs [44]. We employed GPT-4o to generate 350 CVs across six job categories (dentist, nurse, photographer, software engineer, accountant, and teacher), resulting in a total of 2,100 CVs. To mitigate the risk of generating artificial or inaccurate CVs, we instructed GPT-4o to base these CVs on real-world labeled biographies obtained from the BIOS dataset [45]. For the job postings, we utilized 1,358 samples of job ads from a UK job board. They are labeled and from the same categories as the CVs. Therefore, the ground truth is established through the match in job category between CV and job ad. The data was split into training (80%) and validation/test (20%) sets. Our model was trained on a binary classification task to determine whether pairs of CVs and job ads matched. For each positive training instance, we included four negative samples to balance the training process. To track the existence of the FPE, we evaluated the model on the test set containing 272 job ads and 336 unique applicants. We predict the match between pairs of candidates and job ads with relevance score, and counted 13,607 true positive (TP) and 1,625 false negative (FN) predictions, where TP means the candidate is suitable for the job ad and the model predicts it correctly, while FN means the candidate is suited for the job ad but the model predicts they are unsuited.

We conducted two experiments to study the FPE, in particular, to which extent it could apply to algorithmic ranking or matching. In the first experiment, we considered pairs of job ads and candidates; the former as p , the latter as q . TP samples can therefore be considered as fulfilling *if p then q*; FN samples as fulfilling *if p then not q*. As stimuli we considered the presence (or absence) of adjectives in job ads, and investigated to which extent the FPE would extend to the learned ranking model, simulating the situations that the classifier “sees” (or does not “see”) these adjectives during decision making. To this end, we removed a percentage of randomly selected adjectives from the job ads constituting TP samples. As shown in Figure 1, we found that as more adjectives were removed from job ads, more CVs that were TPs (q) became FNs (*not q*). This points to the crucial role of adjectives in RSs’ decision making of applicants’ suitability for a job even though, for instance, “a passionate dentist” should objectively be treated the same as “a dentist”. Yet the recommendation model relies on the adjectives to determine suitability of candidates.

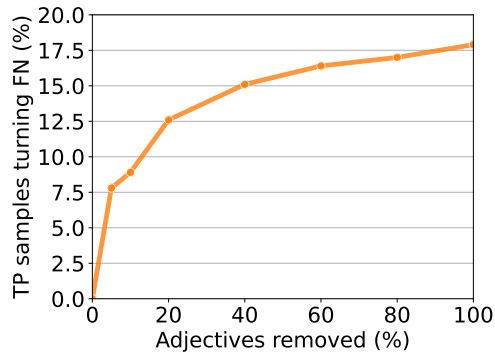


Figure 1: The effect of removing adjectives from job ads on the relevance predictions of candidate-job pairs.

Table 1

Examples of adjectives that exist in each unique set of job ads determined by Textblob [46].

Group	Adjectives
Low Recall	small, referral, sexual, steady, ...
High Recall	new, full, other, good, professional, ...
Unique set	technical, annual, innovative, complex, ...

In a second experiment, we explored leveraging missing adjectives to enhance the decision making of the model. Within the formal logical framework, this can be interpreted as slightly modifying p (by replacing words in the job ad), and subsequently investigating whether *if p then not q* becomes *if p then q* after this modification. To modify p , we first calculated recall for each job ad with respect to all candidates (CVs) in the test set. Subsequently, we grouped the job ads into low- and high-recall categories. We then identified a unique set of adjectives present in high-recall job ads but absent in low-recall job ads. Please refer to Table 1 for some examples. Finally, we randomly selected adjectives from the FN samples and replaced them with adjectives that occurred exclusively in the high-recall job ads (unique set). Overall, we observed a notable improvement in the average score of these modified FN samples, which increased from 0.046 to 0.152. In detail, we observed that 52.0% of the FN samples showed an improvement in their relevance score and 12.9% were reclassified as TPs. In contrast, 39.4% decreased their score and remained FNs while the remaining 8.6% of samples did not change their score. This observation hints at the potential advantage of using missing information to improve the utility of candidate ranking systems.

These findings point at the significance of the FPE, not only from a human perspective, but also evidenced by our algorithmic interpretation of the FPE in a learned ranking model. Our experiments highlight the influence of missing information, such as adjectives, on the accuracy of content-based candidate-job matching systems.

Potential for Exploitation: *The impact of the (non-)existence of certain adjectives in job ads on the accuracy of candidate-job matching is particularly noteworthy when it comes to transparency. The FPE could be mitigated, for instance, by giving recruiters direct feedback on how the pool of applicants changes when they adjust the wording of their job ad. Likewise, a similar transparency mechanism could help job seekers identify salient words in their CVs, or even investigate counterfactual recommendations if they alter aspects like their gender or work experience.*

3.2. Ikea Effect

In the context of RSs and streaming platforms, the Ikea effect could be interpreted as a user’s predisposition towards items and item collections they feel invested in, such as items they conducted research on or discovered themselves, collections they helped compile. To probe the effect in the music domain, we conduct a user study, striving to answer the following research question: Do music listeners prefer

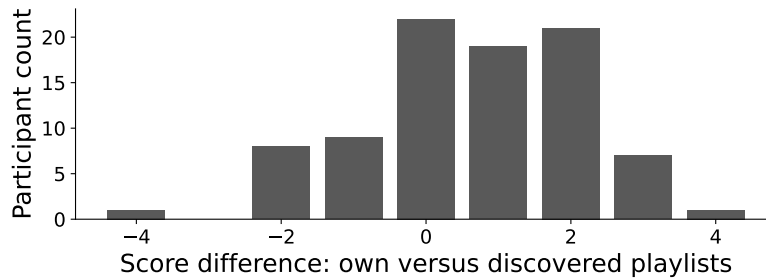


Figure 2: Distribution of the consumption frequency difference between own and other playlists. Positive values show preference towards own playlists.

playlists they contributed to (*own*) over playlists created without their participation (*other*)? We ran the study on the Prolific³ platform for 100 participants from the United States who indicated themselves as users of one or more music streaming services. The study requires participants to complete four statements by choosing one option from a Likert-5 scale ranging from “Never (1)” to “Very often (5)”. The statements are presented in the same order to all participants: S1 “I **create or edit** music collections...”,⁴ S2 “I **play** music collections (created by me or someone else)”, S3 “I play music collections **I created or helped create myself**”, S4 “I play music collections **created by someone else** (e.g., discovered on the internet or shared by someone)”. Out of 96 respondents who submitted valid attempts, 88 indicated that they create or edit as well as consume playlists more often than “Never” (S1 and S2). Out of the 88 participants creating or editing playlists, 48 indicated that they consume own playlists more often than other (comparing responses to S3 and S4), and 18 participants indicated the opposite. The distribution of the difference in consumption frequency between own and other playlists (the S4 score subtracted from the S3 score for each user) is shown in Figure 2. The corresponding mean of these (S3 score – S4 score) is 0.65 and the standard deviation is 1.52. For users creating playlists, we also analyze Spearman’s correlation between responses. A significant correlation between S1 and S3 (0.75, $p < 0.001$) shows that the more time users put into creating playlists the more time they spend listening to their own playlists. On the other hand, we observe no significant correlation between S1 and S4. As an additional finding, we also note the significant correlation between S2 and S3 (0.66, $p < 0.001$). It shows that respondents spending more time listening to playlists in general tend to listen to the playlists they contributed to more often. We interpret this as active listeners having a clearer formed taste and therefore preferring their own playlists to cater to it. This particular finding, while not a direct evidence of the Ikea effect, may indicate a higher degree of investment active listeners feel towards their chosen music. We observe no significant correlation between S2 and S4.

From our observations we conclude that, in the given sample, users tend to interact more with playlists they invested effort in, which we interpret as a variant of the Ikea effect. A larger study would be required to evaluate the prevalence of the effect globally and investigate factors contributing to the users’ feeling of being invested into an item.

Potential for Exploitation: *Acknowledging the concept of user investment into an item or a collection within a RS could help improve user experience with it. For instance, in the scenario of sequential recommendation, items present in the user’s playlists (the user put effort into picking and assigning them) can be seen as particularly valuable and serve as anchor points to retain user engagement within the current listening session. Additionally, recommendation explanations based on items and collections the user feels invested in could help improve user trust (Ikea effect combined with halo effect).*

³<https://www.prolific.com>

⁴The participants were instructed to interpret “music collection” as playlist, track compilation, mixtape or similar, created by a human or automatically.

3.3. Cultural Homophily

Homophily is often defined as the tendency of individuals to associate with similar others [20], e.g., sharing same interests, language, or culture. In the context of RSs, it could be interpreted through the interaction between item consumers and item creators. Users preferring items produced by creators sharing their cultural background would be an example of such interpretation [21]. In the following, we present the results of an empirical study showing indications of cultural homophily in the domain of music recommendation. We put forward two research questions: (1) Do users tend to prefer music tracks produced by artists from their own country? and (2) Can RSs foster or counteract homophily in the setting of music recommendation? We answer the first question by analyzing the listening behavior of users from different countries. We answer the second question by conducting a feedback loop simulation, comparing recommendations provided by a RS at various steps of the feedback loop with actual user consumption behavior.

Following [23], we select a 5-core filtered sample from years 2018-2019 of the LFM-2b dataset [47] containing 99,897 items (songs) selected uniformly at random and 2,287,732 interactions triggered by 11,776 users of the music platform Last.fm.⁵ We enrich the sample with information about the country of each artist crawled from MusicBrainz.⁶ We conduct a feedback loop simulation, following [48]. Feedback loop is a setting where a RS (ranking model) is iteratively updated (retrained) using interactions triggered by the users interacting with previous states of the system. In a way the recommender is trained on the items it partially recommended itself. In our offline setting, we therefore simulate user interactions within the system. At each of the 20 iterations, the data used for training on the previous step is enriched with one recommendation per user, selected from their personalized recommendations, with higher probability for higher ranking recommendations, and then used for training at the next iteration. We use MultVAE [49] as recommendation algorithm.

Table 2 shows the proportions of domestic music in the data sample, user consumption behavior, and user recommendations for countries with the top 10 largest number of tracks in the data sample. The value 0.397 for the *US, base* shows that about 40% of unique tracks in the data sample are produced by artists from the *US*. This column serves as a baseline, showing what the proportion of domestic consumption would be if every user were to interact with tracks chosen at random. The value 0.626 for the *US, Con* shows that over 60% of actual listening attention of US users is directed towards tracks produced by US artists. Respectively *US, Rec₂₀* of 0.595 means that among all recommendations to US users at iteration 20, slightly less than 60% of tracks were produced by US artists.⁷

We answer the first question comparing the columns *base* and *Con*, noticing that for all presented countries the proportion of actual consumption of domestic music is higher than the baseline. This shows that on average users have higher interest in music originating from their country than a random choice would warrant. We also notice that Finnish (*FI*) users demonstrate a higher level of interest in their domestic music than users from countries with comparable track supply, e.g., Australia (*AU*) and Brazil (*BR*), by inspecting column *Con/base* showing domestic consumption relative to the available domestic supply.

As for the second question, focusing on the first iteration of the feedback loop (*Rec₁*), we find that users from the US get recommended a proportion of their domestic music similar to their actual consumption (*Con*) while users from Germany (*DE*), Brazil (*BR*), and Russia (*RU*) are slightly over-served with their domestic music. In contrast, users from other countries are under-served with their domestic music, most prominently users from Australia (*AU*), whose recommendations fail to reach the level of the baseline. At iteration 20 of the loop (*Rec₂₀*) the proportion of recommended domestic music increases for some under-served countries, such as the United Kingdom (*UK*) and Canada (*CA*). However, for other countries such as Sweden (*SE*) and Finland (*FI*) the proportion of recommended domestic music decreases further away from their actual consumption level.

⁵<https://www.last.fm>

⁶<https://www.musicbrainz.org>

⁷Note that in the case of columns *Con* and *Rec_c*, a track is counted multiple times if it is consumed by, or recommended to, multiple users (also taken into account for normalization).

Table 2

Proportions of domestic music among all available tracks (*base*), among consumed tracks by users from the country (*Con*), among recommended tracks (*Rec₁*), and among recommended tracks after 20 iterations of the feedback loop simulation (*Rec₂₀*), for the top 10 countries in the dataset. Bold values indicate the highest, underlined lowest values.

	<i>base</i>	<i>Con</i>	<i>Con/base</i>	<i>Rec₁</i>	<i>Rec₁/base</i>	<i>Rec₂₀</i>	<i>Rec₂₀/base</i>
<i>US</i>	0.397	0.626	<u>1.578</u>	0.629	1.587	0.595	1.501
<i>UK</i>	0.155	0.266	1.713	0.227	1.458	0.232	1.495
<i>DE</i>	0.068	0.169	2.481	0.176	2.590	0.166	2.439
<i>SE</i>	0.045	0.159	3.519	0.102	2.266	0.088	1.948
<i>CA</i>	0.038	0.083	2.202	0.030	0.797	0.041	<u>1.091</u>
<i>FR</i>	0.028	0.091	3.232	0.039	1.377	0.041	1.447
<i>AU</i>	0.023	0.077	3.289	<u>0.017</u>	<u>0.728</u>	<u>0.026</u>	1.103
<i>FI</i>	0.023	0.170	7.536	0.166	7.325	0.132	5.820
<i>BR</i>	0.022	0.141	6.288	0.187	8.347	0.150	6.714
<i>RU</i>	<u>0.019</u>	<u>0.073</u>	3.870	0.081	4.262	0.066	3.515

The results of this preliminary study show that users tend to be interested in music originating from their own country, potentially indicating cultural homophily. However, the degree of interest varies between countries, e.g., Finnish users consume more domestic music than Australian users while having a comparable domestic supply. RSs do not necessarily represent this interest in their output and could foster it for some countries while counteracting it for others.

Potential for Exploitation: *Formalizing and leveraging corresponding homophily models as an additional indicator of user taste could be useful to strike a better balance in terms of diversification of recommendations, or to pursue calibration between user profiles and recommendations in terms of country, which is not always achievable in contemporary RSs [50].*

4. Conclusion, Limitations, and Future Work

Cognitive biases have recently gained recognition in the wider field of machine learning, notably also in learning to rank tasks. While the recommender systems community has addressed biases like confirmation bias and position bias to some degree, other cognitive biases have remained largely unexplored. Addressing this gap, we demonstrated the presence of several cognitive biases within the recommendation ecosystem, including feature-positive effect, Ikea effect, and cultural homophily. We also briefly outlined potential positive and negative aspects of these biases in the recommendation process.

Limitations: The reported experiments represent only first steps towards obtaining a better understanding of the studied cognitive biases in the context of RSs. Our study covers three biases and the two recommendation domains of recruitment and music. While this may limit the generalizability of the findings to other domains, we believe future studies could build upon the ones we conducted here. Another limitation of these preliminary studies is that the experiments assume traditional recommendation setups (e.g., matching jobs with CVs for the feature-positive effect and top-k recommendation for homophily), neglecting the presumably strong influence the RS’s user interface has on the manifestation or amplification of cognitive biases. Also, some of the presented results may be influenced by the characteristics of underlying datasets. For example, the LFM-2b dataset may be representative of users of the Last.fm service, but the country distributions of users and artists may not necessarily be representative of the population at large. Concerning cultural homophily, in the absence of better cultural indicators for both users and items/artists we resort to their country of origin. This limits generalizability of our findings because a song produced in one country does not need to be written in the (official) language of the country’s citizens, nor represent their culture. In addition, cultural identities can extend beyond

country borders. Future work therefore calls for investigating additional cultural aspects, for instance, based on religion, interwoven history, or Hofstede's cultural dimensions [51].

Future Work: The next steps in our research agenda include devising rigorous, formalized models of cognitive biases, creating a framework to precisely describe how they influence recommendation outcomes, and developing prototype RSs that leverage the positive effects of cognitive biases while mitigating the negative ones for recommendation. Focusing on the three biases studied here, concrete future directions may include investigating the dependence of the Ikea effect on different functionalities offered by different streaming platforms and RSs which alters the ways users interact with them. Given the varying prevalence of certain RS platforms between countries, we expect notable differences. The next steps in the study of cultural homophily could include disentangling homophily from other factors, such as availability bias, as well as studying the relation between homophily and diversity in recommendations.

In conclusion, we strongly believe that RS researchers and practitioners should be equipped with high sensitivity for the existence of cognitive biases and knowledge of how they may impact different stages in the recommendation process. Furthermore, forming an understanding of where cognitive biases originate from, how they are intertwined, and how they influence different stakeholders will be vital to fully comprehend their role in the recommendation ecosystem. We advocate for a holistic discussion of both negative and positive effects of cognitive biases, and for new recommendation approaches that mitigate the former while exploiting the latter. In our opinion, this will be highly beneficial for the RS community as a whole, and may even mark the starting point for a new generation of cognitive-bias-aware RSs.

Acknowledgments

This research was funded in whole or in part by the Austrian Science Fund (FWF): <https://doi.org/10.55776/P33526>, <https://doi.org/10.55776/DFH23>, <https://doi.org/10.55776/COE12>, <https://doi.org/10.55776/P36413>. Furthermore, we thank Stefan Brandl, Gustavo Escobedo, and Mohammad Lotfi for their contributions to initial experiments.

References

- [1] D. Kahneman, *Thinking, Fast and Slow*, Penguin Books, 2017.
- [2] R. Dobelli, *The Art of Thinking Clearly: Better Thinking, Better Decisions*, Hachette, UK, 2013.
- [3] J. L. Eberhardt, *Biased: Uncovering the Hidden Prejudice That Shapes What We See, Think, and Do*, Penguin Books, 2020.
- [4] E. Bonilla-Silva, *Racism Without Racists: Color-blind Racism and the Persistence of Racial Inequality in the United States*, Rowman & Littlefield Publishers, 2006.
- [5] A. Tversky, D. Kahneman, Judgment Under Uncertainty: Heuristics and Biases: Biases in Judgments Reveal Some Heuristics of Thinking Under Uncertainty, *Science* 185 (1974) 1124–1131.
- [6] G. Gigerenzer, W. Gaissmaier, Heuristic Decision Making, *Annual Review of Psychology* 62 (2011) 451–482.
- [7] S. Fabi, T. Hagedorff, Why We Need Biased AI - How Including Cognitive and Ethical Machine Biases Can Enhance AI Systems, *CoRR* abs/2203.09911 (2022). [arXiv:2203.09911](https://arxiv.org/abs/2203.09911).
- [8] L. Azzopardi, Cognitive biases in search: A review and reflection of cognitive biases in information retrieval, in: *Proceedings of CHIIR, 2021*, p. 27–37.
- [9] J. Liu, L. Azzopardi, Search Under Uncertainty: Cognitive Biases and Heuristics: A Tutorial on Testing, Mitigating and Accounting for Cognitive Biases in Search Experiments, in: *Proceedings of SIGIR, 2024*, p. 3013–3016.

- [10] G. Gomroki, H. B. and.. Rahmatolloah Fattahi, J. S. Fadardi, Identifying Effective Cognitive Biases in Information Retrieval, *Journal of Information Science* 49 (2023) 348–358.
- [11] K. Lerman, T. Hogg, Leveraging Position Bias to Improve Peer Recommendation, *PLOS ONE* 9 (2014) 1–8.
- [12] E. C. Teppan, M. Zanker, Decision Biases in Recommender Systems, *Journal of Internet Commerce* 14 (2015) 255–275.
- [13] E. Lex, D. Kowald, P. Seitlinger, T. N. T. Tran, A. Felfernig, M. Schedl, Psychology-informed Recommender Systems, *Foundations and Trends in Information Retrieval* 15 (2021) 134–242.
- [14] M. Schedl, V. W. Anelli, E. Lex, Technical and Regulatory Perspectives on Information Retrieval and Recommender Systems: Fairness, Transparency, and Privacy, Springer Cham, 2024.
- [15] S. T. Allison, D. M. Messick, The Feature-positive Effect, Attitude Strength, and Degree of Perceived Consensus, *Personality and Social Psychology Bulletin* 14 (1988) 231–241.
- [16] E. Rassin, Individual Differences in the Susceptibility to the Feature-Positive Effect, *Psychological Test Adaptation and Development* (2023).
- [17] M. D. Ekstrand, A. Das, R. Burke, F. Diaz, Fairness in Information Access Systems, *Foundations and Trends in Information Retrieval* 16 (2022) 1–177.
- [18] M. I. Norton, D. Mochon, D. Ariely, The IKEA Effect: When Labor Leads to Love, *Journal of Consumer Psychology* 22 (2012) 453–460.
- [19] L. E. Marsh, P. Kanngiesser, B. Hood, When and How Does Labour Lead to Love? The Ontogeny and Mechanisms of the IKEA Effect, *Cognition* 170 (2018) 245–253.
- [20] N. Ferguson, The False Prophecy of Hyperconnection: How to Survive the Networked Age, *Foreign Affairs* 96 (2017) 68–79.
- [21] N. P. Mark, Culture and Competition: Homophily and Distancing Explanations for Cultural Niches, *American Sociological Review* 68 (2003) 319–345.
- [22] Y. Jiang, D. I. Bolnick, M. Kirkpatrick, Assortative Mating in Animals, *The American Naturalist* 181 (2013) E125–E138.
- [23] O. Lesota, E. Parada-Cabaleiro, S. Brandl, E. Lex, N. Rekabsaz, M. Schedl, Traces of Globalization in Online Music Consumption Patterns and Results of Recommendation Algorithms, in: *Proceedings of ISMIR, 2022*, pp. 291–297.
- [24] M. Mather, L. L. Carstensen, Aging and Motivated Cognition: The Positivity Effect in Attention and Memory, *Trends in Cognitive Sciences* 9 (2005) 496–502.
- [25] A. R. Sutin, A. Terracciano, Y. Milaneschi, Y. An, L. Ferrucci, A. B. Zonderman, The Effect of Birth Cohort on Well-Being: The Legacy of Economic Hard Times, *Psychological Science* 24 (2013) 379–385.
- [26] T. R. Mitchell, L. Thompson, E. Peterson, R. Cronk, Temporal Adjustments in the Evaluation of Events: The “Rosy View”, *Journal of Experimental Social Psychology* 33 (1997) 421–448.
- [27] E. Parada-Cabaleiro, M. Mayerl, S. Brandl, M. Skowron, M. Schedl, E. Lex, E. Zangerle, Song Lyrics Have Become Simpler and More Repetitive Over the Last Five Decades, *Scientific Reports* 14 (2024) 5531.
- [28] X. J. Yang, C. D. Wickens, K. Hölttä-Otto, How Users Adjust Trust in Automation: Contrast Effect and Hindsight Bias, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 60 (2016) 196–200.
- [29] G. B. Chapman, E. J. Johnson, Incorporating the Irrelevant: Anchors in Judgments of Belief and Value, *Heuristics and Biases: The Psychology of Intuitive Judgment* (2002).
- [30] J. W. Payne, J. R. Bettman, E. J. Johnson, *The Adaptive Decision Maker*, Cambridge University Press, 1993.
- [31] E. C. Teppan, A. Felfernig, Calculating Decoy Items in Utility-based Recommendation, in: *Proceedings of IEA/AIE, Springer, 2009*, pp. 183–192.
- [32] G. Adomavicius, J. Bockstedt, S. Curley, J. Zhang, Recommender Systems, Consumer Preferences, and Anchoring Effects, in: *Proceedings of the Workshop on Human Decision Making in Recommender Systems, 2011*, pp. 35–42.
- [33] Y. Zheng, C. Gao, X. Li, X. He, Y. Li, D. Jin, Disentangling User Interest and Conformity for

- Recommendation with Causal Embedding, in: Proceedings of The Web Conference, 2021, pp. 2980–2991.
- [34] C. Ma, Y. Ren, P. Castells, M. Sanderson, Temporal Conformity-aware Hawkes Graph Network for Recommendations, in: Proceedings of The Web Conference, 2024, pp. 3185–3194.
- [35] N. Craswell, O. Zoeter, M. J. Taylor, B. Ramsey, An Experimental Comparison of Click Position-Bias Models, in: Proceedings of WSDM, 2008, pp. 87–94.
- [36] A. Felfernig, G. Friedrich, B. Gula, M. Hitz, T. Kruggel, G. Leitner, R. Melcher, D. Riepan, S. Strauss, E. Teppan, O. Vitouch, Persuasive Recommendation: Serial Position Effects in Knowledge-Based Recommender Systems, in: Proceedings of PERSUASIVE, volume 4744 of *Lecture Notes in Computer Science*, Springer, 2007, pp. 283–294.
- [37] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, Y. Qin, An Integrated Theory of the Mind, *Psychological Review* 111 (2004) 1036.
- [38] M. Reiter-Haas, E. Parada-Cabaleiro, M. Schedl, E. Motamedi, M. Tkalcic, E. Lex, Predicting Music Relistening Behavior Using the ACT-R Framework, in: Proceedings of RecSys, 2021, pp. 702–707.
- [39] M. Moscati, C. Wallmann, M. Reiter-Haas, D. Kowald, E. Lex, M. Schedl, Integrating the ACT-R Framework with Collaborative Filtering for Explainable Sequential Music Recommendation, in: Proceedings of RecSys, 2023, pp. 840–847.
- [40] R. E. Nisbett, T. D. Wilson, The Halo Effect: Evidence for Unconscious Alteration of Judgments, *Journal of Personality and Social psychology* 35 (1977) 250.
- [41] E. L. Thorndike, A Constant Error in Psychological Ratings, *Journal of Applied Psychology* 4 (1920) 25–29.
- [42] C. Batres, V. Shiramizu, Examining the “Attractiveness Halo Effect” Across Cultures, *Current Psychology* 42 (2023) 25515–25519.
- [43] H. S. Jenkins, The Development of Stimulus Control Through Differential Reinforcement, *Fundamental Issues in Associative Learning* (1969).
- [44] D. Kumar, T. Grosz, E. Greif, N. Rekasaz, M. Schedl, Identifying Words in Job Advertisements Responsible for Gender Bias in Candidate Ranking Systems via Counterfactual Learning, in: Proceedings of RecSys in HR, volume 3490, CEUR-WS.org, 2023.
- [45] M. De-Arteaga, A. Romanov, H. Wallach, J. Chayes, C. Borgs, A. Chouldechova, S. Geyik, K. Kenthapadi, A. T. Kalai, Bias in Bios: A Case Study of Semantic Representation Bias in a High-Stakes Setting, in: Proceedings of FAccT, 2019, pp. 120–128.
- [46] S. Loria, Textblob Documentation, Release 0.15.2 (2018).
- [47] M. Schedl, S. Brandl, O. Lesota, E. Parada-Cabaleiro, D. Penz, N. Rekasaz, LFM-2b: A Dataset of Enriched Music Listening Events for Recommender Systems Research and Fairness Analysis, in: Proceedings of CHIIR, 2022, pp. 337–341.
- [48] M. Mansoury, H. Abdollahpouri, M. Pechenizkiy, B. Mobasher, R. Burke, Feedback Loop and Bias Amplification in Recommender Systems, in: Proceedings of CIKM, 2020, pp. 2145–2148.
- [49] J. Xu, Y. Ren, H. Tang, X. Pu, X. Zhu, M. Zeng, L. He, Multi-VAE: Learning Disentangled View-common and View-peculiar Visual Representations for Multi-view Clustering, in: Proceedings of ICCV, 2021, pp. 9214–9223.
- [50] O. Lesota, J. Geiger, M. Walder, D. Kowald, M. Schedl, Oh, Behave! Country Representation Dynamics Created by Feedback Loops in Music Recommender Systems, in: Proceedings of RecSys, 2024.
- [51] G. Hofstede, Dimensionalizing Cultures: The Hofstede Model in Context, *Online Readings in Psychology and Culture* 2 (2011).