Argumentative explanations for recommendations - Effect of display style and profile transparency

Diana C. Hernandez-Bocanegra
University of Duisburg-Essen
Duisburg, Germany
diana.hernandez-bocanegra@uni-due.de

Jürgen Ziegler
University of Duisburg-Essen
Duisburg, Germany
juergen.ziegler@uni-due.de

ABSTRACT

Providing explanations based on user reviews in recommender systems may increase users' perception of transparency. However, little is known about how these explanations should be presented to users in order to increase both their understanding and acceptance. We present in this paper a user study to investigate the effect of different display styles (visual and text only) on the perception of review-based explanations for recommended hotels. Additionally, we also aim to test the differences in users' perception when providing information about their own profiles, in addition to a summarized view on the opinions of other users about the recommended hotel. Our results suggest that the perception of explanations regarding these aspects may vary depending on user characteristics, such as decision-making styles or social awareness.

CCS CONCEPTS

• Information systems → Recommender systems; **• Human-centered computing → User studies**.

KEYWORDS

Recommender systems, user study, explanations

1 INTRODUCTION

Nowadays, recommender systems (RS) are still perceived as black boxes by users, where little can be done to obtain the reasons that justify the recommendations. Providing explanations of the rationale behind a recommendation can bring several benefits to RS. In particular, explanations may serve the aims of transparency (the system explains how it works) and effectiveness (user can make good decisions) [21], among other explanatory aims. Among the most popular approaches are the feature-based explanations, which can provide rationale for content-based methods by providing users with item features that match their preferences (e.g. [20]), as well as explanations based on relevant users or items, which provide rationale for collaborative filtering methods (e.g. [7]). More recently, there has been increased interest in the use of user reviews in explanation methods, given the richness of information reported on diverse aspects, which cannot be deduced from the overall item ratings. The above represents a potential for the generation of argumentative explanations, which seek to provide more robust statements on both positive and negative aspects reported by users in order to support an item recommendation, compared to shallow sentences like Amazon's "Customers who bought … also bought...".

Currently, while one of the most used approaches to evaluate the quality of review based explanations is the use of offline evaluation metrics (e.g. BLEU [16] and ROUGE [13] in the case of natural language generated statements), a more comprehensive empirical assessment of users' perception is still needed, in order to answer a relevant question that remains open: how to present argumentative explanations in an comprehensible manner, in order to meet the explanatory aims of transparency and effectiveness? Previously, and with the aim of deepening this matter, [8] compared different types of review-based argumentative explanations in the hotel domain, and found that users perceived a higher explanation quality when an aggregated view of positive and negative opinions using percentages was provided, compared to a summary of opinions without providing any percentage; furthermore, a greater perceived transparency was reported for explanations with the aggregated view using percentages of opinions, compared to explanations that only provided a useful review. Thus, showing a consolidated view of other users' opinions using percentages seems to be an effective way of providing explanations in systems that assist the evaluation of items such as hotels. However, it remains unclear what is the most proper way to display statistical information based on user reviews as part of the explanations. Here, for example, users with lower visual abilities might benefit less from a display based on images or graphics [10, 18], compared to a presentation using only text.

Depending on the method used to generate the recommendations, the explanations may reflect either the properties of the items or the preferences of the users, which were used to generate the recommendations. In terms of displaying user preferences as part of explanations, a target user might be benefited from knowing which of her/his performed interactions with the system are having an effect on a current recommendation, as pointed out by [7]. Although providing a view on user profiles in content-based or item-based collaborative filtering methods might be considered as an usual practice (e.g. [7, 23]), information on users' profile is often omitted in review-based explanations, and used only implicitly, e.g. to filter and sort lists of relevant features, as in [15]. Thus, a question that remains open is to what extent providing information on user preferences influences the perception of review-based explanations by users.

Consequently, we set out to answer the following research questions in regard to different explanatory aims (explanation quality,

MuC'20 Workshops, Magdeburg, Deutschland

https://doi.org/10.18420/muc2020-ws111-338

transparency, effectiveness, efficiency, and trust) and review-based explanations:

RQ1: Does the display *style* of explanation (using charts or only text) influence the perception of the variables of interest?

RQ2: Does including or not the information about *user preferences* influence the perception of the variables of interest?

Similarly to [8], we also aimed to test the effect that user characteristics may have on the perception of the explanations, in particular regarding decision making style (rational and intuitive) [6] and the ability of the user to take into account the views of others (social awareness) [5]. Additionally, we also aimed to test the influence that visual familiarity may have on explanations perception, for which we used the items proposed by [12]. Consequently:

RQ3: Do individual differences in visual familiarity, social awareness or decision making styles influence the perception of our proposed explanations design?

In order to address these questions, we conducted a user study to test the perception of explanations based on user opinions in the hotel domain, given different display styles and whether or not user profile information is shown. Explanation design is elaborated in detail in section 3. Additionally, we based our approach on the explicit factor model (EFM) proposed by [27], a matrix factorization based model that seeks to align implicit features with explicit features extracted from user reviews, to generate explanations.

2 RELATED WORK

Review-based explanation methods leverage the richness of detailed information on positive and negative item aspects reported by users in their reviews, information that is usually hard to grasp from overall item ratings. By using these methods, the following types of explanation can be generated: **1)** An abstract summarization of review findings, i.e. statements generated in natural language representing a summarized version of the original content extracted from reviews, e.g. [3], who proposed a method based on natural language generation (NLG) techniques. **2)** A selection of helpful reviews that might be relevant to users, as proposed by [2], who uses a deep learning model and word embeddings to jointly learn user preferences and item properties, and an attention mechanism to detect features that are of most interest to the target user. **3)** A summarized view of pros and cons on specific item aspects reported by other users. Here, topic modelling and aspect-based sentiment analysis are usually used to detect the sentiment polarity towards item aspects or features addressed in reviews, as in [4, 26, 27]. Subsequently, such information can be integrated to RS algorithms like matrix or tensor factorization, as in [1, 25, 27] in order to generate both recommendation and aspect-based explanations.

In this paper we focus on the third approach. Here, explanations proposed are usually presented to users either providing 1) a summary of the positive and negative opinions on different aspects e.g. using bar charts as in [15] or word clouds and radar charts as in [26], or 2) sentences generated using templates, e.g. "You might be interested in [feature], on which this product performs well". Although bar charts reflecting positive and negative views might be perceived as more informative than brief template-based textual explanations, or even easier to interpret than challenging radar charts, it remains unclear whether there is a difference in the perception of review-based explanations when they are presented using visual representations (e.g. charts) as opposed to only text. In this regard, [12] proposed a series of explanations based on a hybrid RS in the music domain, and tested, among others, the influence that the presentation format could have on users' perception. In this case, the authors found that textual explanations were perceived as more persuasive than the explanations provided using a visual format; however, users with greater visual familiarity perceived one of the visual format explanations more positively (a Venn diagram).

In regard to the types of information provided, while it is common to present an overview of user preferences and their relation to item properties in content-based and item-based explanation methods (e.g. [7, 23]), little is known about the effect of presenting the user profile as part of review-based explanations. In fact, in the examples provided by [15], the user preferences are only implicitly represented (i.e. aspects are sorted by user importance in [15], but this is not communicated to users), unlike [26], which provides them explicitly. However, as no user evaluation was applied in the latter case, it remains unclear to what extent providing such information influences the perception of this kind of explanation by the target users.

Finally, in order to address our above mentioned research questions, we implemented the explicit factor model (EFM) proposed by [27], and extended their work by proposing and testing the perception of users of a set of argumentative explanations, that seek to provide more robust reasons behind recommendations in comparison to the brief template-base explanations proposed in the original EFM work. The EFM model uses, in addition to the user-item ratings matrix, two additional matrices: an item quality matrix (including number of negative and positive comments per aspect) and an user preference matrix (times that a user mentioned an aspect in her/his reviews); by using these matrices, it is possible to generate explanations involving both user preferences and item profiles, which can be represented both in visual and only text styles.

3 EXPLANATION DESIGN

In the context of recommendation systems, review-based argumentative explanations could be understood as a set of propositions, summarizing positive points reported by other users on specific aspects, that support the claim that an article can be recommended to a user. In this respect, information extracted from user reviews could be consolidated and provided as propositions, which would constitute the *backing* component according to the argumentative scheme proposed by Toulmin [22], while the conclusion (the item is recommended) constitute itself a *claim*. While this could be considered a 'shallow' structure, compared to the complete Toulmin argument scheme (which involves additional components, like rebuttal or refutation), it resembles explanation schemes based on deductive arguments, such as those widely used in the scientific field (i.e. a set of explanatory propositions is logically followed by an explanatory target, as discussed by [19]), or even more particularly, explanation schemes in RS such as the one used by [**?**], who provides brief sentences - two facts and a claim - as explanations for content-based recommendations of hiking routes, energy and mobile phone plans.

In consequence, our explanation design seeks to represent an argumentative structure, while reflecting in turn the arguments provided by other users in their reviews, in a consolidated manner. Therefore, our proposed scheme consists of a claim ("We recommend this hotel") and the propositions that support such claim, connected with the conjunction "because". We propose to provide the following pieces of information within proposition statements:

1. Item quality: A summary of comments reported by previous hotel guests for different aspects, as well as what percentage were positive and negative.

2. User preferences: what are the most important item aspects to the target user. In this regard, we aimed to make the user's own profile transparent, by showing the user's inferred importance of each aspect, together with the opinions of other users about the aspect (as shown in the examples included in figures 1a and 1c), in order to facilitate a direct comparison of the points of view of others and their alignment with their own preferences.

3. Statements that inform how the user preferences and item quality are extracted (e.g,"based on how often you mentioned these features in your own comments before"). We believe that providing this information, in addition to the information listed above, could increase the perception of trust by users, while decreasing the perception that they are interacting with a black box.

While arguments are usually associated with oral or written speech, arguments can also be communicated using visual representations (e.g. graphics or images). In this regard, according to [9], visual arguments (a combination of visual and verbal communication) may, in addition to representing propositional content, have a greater "rhetorical power potential" than verbal arguments, due (among others) to their greater immediacy.

In consequence, we aimed to test the effect of the two factors: display *style* and display of the *user preferences*. An example of each condition is provided in Figure 1.

'Visual' style: Provides a view of the number of comments per aspect and percentages of positive and negative opinions using bar charts.

'Text' style: Provides the same information used in the visual condition, but instead of using bar charts, presents the information using only text, organized within a table.

Additionally, every display style involves two variations:

User preferences 'yes'. The information about the user preferences is provided.

User preferences 'no'. No information about the user preferences is displayed.

4 EMPIRICAL STUDY

We aimed to compare users' perception of review-based argumentative explanations, given different styles of display (visual or text). In this regard, we hypothesized that users with greater visual abilities would find explanations better when these are provided using visual aids, like a bar chart, in comparison to only textual information (H1), as reported by [12] in the music domain. Additionally, we also aimed to test whether users' positive perception of the RS would increase when information about the user preferences is provided. Here, we hypothesized that users would report a more positive perception of the RS when information about their user preferences is provided (H2), as reported by [7] for explanations based on collaborative filtering methods.

To test the above, we recruited 150 participants (66 female, mean age 39.08 and range between 23 and 73) through Amazon Mechanical Turk. We restricted the execution of the task to workers located in the U.S, with a HIT (Human Intelligence Task) approval rate greater than 95%, and number of HIT's approved greater than 500. We applied a quality check in order to select participants with quality survey responses, i.e. at least 5 of the 6 high priority validation questions were answered correctly, more than 30s were spent on the recommendation step and more than 50s on the evaluation questionnaire. The responses of 46 subjects were discarded due to this quality check (from an initial number of 195 workers), so only the responses of 150 subjects were used for the analysis (statistical power of 85%, α =0.05). Participants were rewarded with $1 plus a bonus up to $0.4 depending on the quality of their response to the question "Why did you choose this hotel?" set at the end of the survey. Time devoted to the task by participants (in minutes): M=8.04, SD= 1.62.

The study follows a 2x2 between-subjects design, and each participant was assigned randomly to one of six conditions that represent the combination of the two factors: display *style* and *user preferences* provided or not. Similarly to [8], we presented participants with a fixed list of 5 hotels that represented the recommendations for a hypothetical hotel search, and a detailed view including an explanation of why every item was recommended. Then, participants were asked to choose the hotel they considered the best, to report their reasons to it, and to rate their perception of both recommender and its explanations. The explanations and recommendations were generated using the EFM algorithm [27] and the dataset of hotels' reviews, ArguAna [24], although they were presented to the participants only through a prototype, i.e. no real system was implemented to allow the interactions.

Given that we could not ensure access to previously written participants' reviews (which is not only important for the optimal functioning of the algorithm, but also constitutes a base to test the condition "user preferences"), we calculated the top 5 of the most important aspects to all users within the dataset, namely: room, price, facilities, location and staff. Then, a random user was chosen from the dataset with those same preferences, and 5 of her top-ranked options according to the EFM algorithm were selected to be presented to participants, alongside their explanations. Additionally, we presented the users with a cover story, in which we told the users to pretend that their most important aspect was the "room" and the "price".

Conditions: We tested two display styles: "visual" (depicts explanations using bar charts) and "text" (using only text). Additionally, we tested the effect of including a view of the user preferences "yes" (when included) "no" (when no information of user preferences is provided). Section 3 provides further details on every display style and user preferences information.

Procedure: First, participants were asked to answer demographic questions and the questionnaire on user characteristics. We indicated in the instructions step that a 5 hotels list reflecting the results of a hypothetical hotels' search would be presented. We asked them to click the "View Details" button for each hotel, and to read carefully the explanations provided in each case (examples of

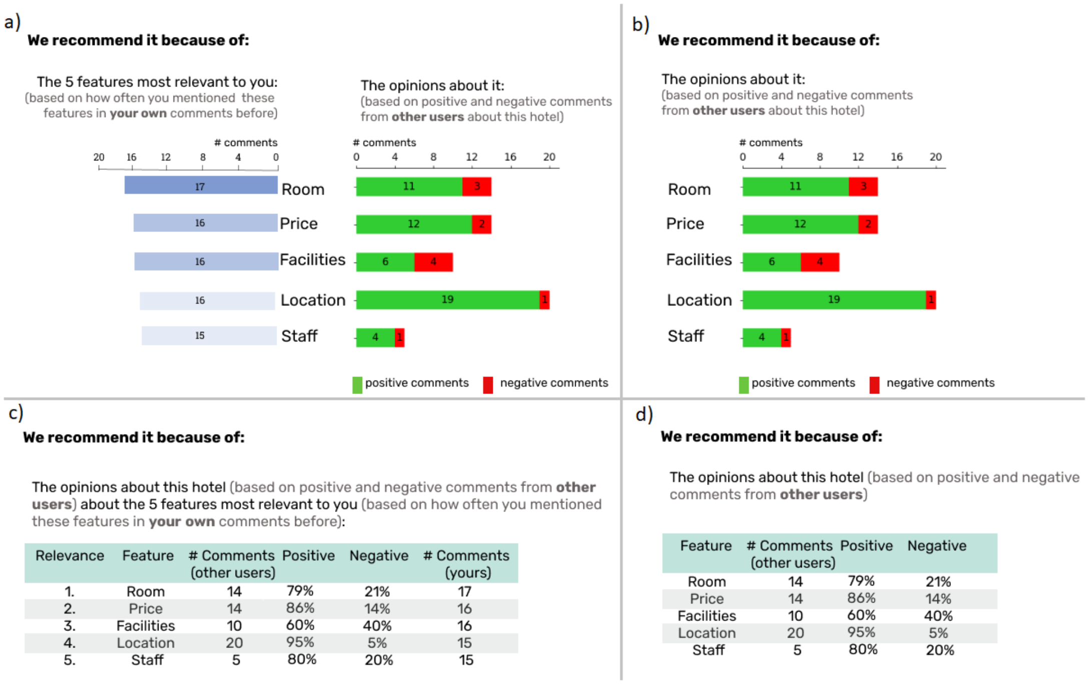

Figure 1: Explanations displayed in empirical study for every experimental condition, for one of the recommended hotels. a) Style 'visual' , user preferences 'yes'. b) Style 'visual' , user preferences 'no'. c) Style 'text' , user preferences 'yes'. d) Style 'text' , user preferences 'no'.

explanations for the different experimental conditions are provided in Figure 1). Additionally, we provided a cover story, as an attempt to establish a common starting point in terms of reasons to travel (a business trip), and the supposedly most interesting aspects for the user (room and facilities).

The list of hotels, their names, photos, prices and locations, as well as their ratings and the numbers of reviews and positive and negative opinions, remained constant to all users. Variations focused only on display style and the presentation of user preferences, depending on the condition to which each participant was assigned. After the interaction with the prototype, subjects were asked to choose the hotel that best suited their purpose, as well as an open question about their reasons for choosing that hotel. Then, subjects answered the evaluation questionnaire. In addition, we included an open-ended question, so that participants could indicate in their own words their general opinion about the explanations provided. We included 11 validation questions to check attentiveness and the effective completion of the task.

Questionnaires:

Evaluation: Similarly to [8], we utilized items from [17] to evaluate the perception of transparency, [11] of effectiveness (user can make good decisions), [14] of efficiency (user can make decisions faster), and [14] of trust. Furthermore, we adapted items from [11] to evaluate explanation quality. All items were measured with a 1-5 Likert-scale (1:Strongly disagree, 5: Strongly agree).

User characteristics: We used all the items of the Rational and Intuitive Decision Styles Scale [6] as well as the scale of the social awareness competency proposed by [5]. Additionally, We used the visual familiarity items as proposed by [12]. All items were measured with a 1-5 Likert-scale (1:Strongly disagree, 5: Strongly agree).

5 RESULTS

Evaluation scores. Evaluation scores were calculated for each variable of interest (quality of the explanation and its objectives: transparency, effectiveness, efficiency and confidence), as the average of the individual values reported for the questionnaire items related to each variable (descriptive results by display style and display of user preferences are reported in Table 1).

Analysis of covariance A MANCOVA analysis was performed, given that our dependent variables are correlated (correlation coefficients included in Table 1). This analysis sought to assess the simultaneous effect of display styles and display of user preferences on the perception of the explanation aims, and to what extent the considered user traits could influence such perception. Here, the dependent variables were the evaluation scores, the independent variables were *style* and *user preferences*, and the covariates were user characteristics. Then, subsequent ANCOVA analyses were performed, in order to test the effect of interactions between independent variables and covariates. The findings are described below.

Multivariate effects:

We found significant multivariate effects for rational $F(6, 138) = 3.75, p < .01$ and intuitive $F(6, 138) = 2.83, p < .05$ decision making style, as well as for social awareness $F(6, 138) = 5.64, p < .001$. No

Table 1: Mean values and standard deviations of perception on explanation aims, per *display style* and *display of user preferences* (n=150); values reported with a 5-Likert scale; high mean values correspond to a positive perception of recommender and its explanations. Pearson correlation matrix, p<0.001 for all correlation coefficients.

Style	Text		Visual		*User Preferences*	Yes		No		*Correlation*	Variable			
Variable	M	SD	M	SD		M	SD	M	SD		1	2	3	4
1. Explanation Quality	3.83	0.65	3.86	0.67		3.88	0.67	3.81	0.654					
2. Transparency	3.80	0.72	3.78	0.80		3.87	0.71	3.72	0.793		0.52	—		
3. Effectiveness	3.88	0.61	3.75	0.75		3.84	0.76	3.79	0.612		0.60	0.49	—	
4. Efficiency	3.96	0.73	3.89	0.92		4.00	0.78	3.86	0.869		0.36	0.39	0.52	—
5. Trust	3.75	0.60	3.89	0.63		3.84	0.65	3.81	0.588		0.66	0.47	0.67	0.58

significant overall effects were found for display style, the display of user preferences, their interaction, nor visual familiarity.

Univariate effects:

Explanation quality:

Neither the display style nor the display of user preferences influences significantly the perception of explanation quality. However, a significant interaction between social awareness and the display of user preferences was found *F(1, 146) = 4.79, p<.05.*, with the "yes" condition having a steeper slope than the "no" condition (showing a positive relationship between social awareness and displaying user preferences), the latter remaining constant regardless of the social awareness score (Figure 2b). Additionally, a main effect of the intuitive decision making style on explanation quality was found *F(1, 146) = 11.35, p<.001*; here, a positive trend was observed between the two variables.

Transparency: No main effects of display style or display of user preferences were found. However, a main effect of social awareness on transparency was observed *F(1, 146) = 8.39, p<.01*; here, we observed a positive trend in the relationship between these two variables (Figure 2d).

Effectiveness:

No main effect of display style or display of user preferences was observed. However, a main effect of rational decision-making style *F(1, 146) = 8.91, p<.01* and social awareness was found *F(1, 146) = 16.10, p<.001*; here, a positive trend is observed between both variables and effectiveness (effect on social awareness in figure 2d). Additionally, although the interaction between display style and rational decision-making style is not significant, *F(2, 146) = 2.82, p=.09*, we observe that the positive trend of the relationship between the "visual" condition and the rational decision-making style was more pronounced than in the "text" condition. (Figure 2c).

Efficiency: A main effect of rational decision-making style *F(1, 146) = 11.63, p<.01* and social awareness was found *F(1, 146) = 8.76, p<.01*; here, a positive trend is observed between both variables and efficiency (effect on social awareness in figure 2d).

Trust: A main effect of rational decision-making style *F(1, 146) = 14.70, p<.001* and social awareness was found *F(1, 146) = 40.49, p<.001*; here, a positive trend is observed between both variables and trust (effect on social awareness in figure 2d).

User characteristics scores. We calculated the scores of the rational (*M* = 4.24, *SD*= 0.56) and the intuitive (*M* = 2.65, *SD*= 1.01) decision making styles, social awareness (*M* = 3.92, *SD*= 0.59) and visual familiarity (*M* = 3.03, *SD*= 1.02) for each individual as the average of the reported values for the items of every scale. Figure 2a shows the distributions of these scores.

6 DISCUSSION

Our results suggest that the display style or the display of user preferences do not influence the perceived quality of the proposed explanations, unless individual differences are taken into account, such as social awareness, i.e the extent to which subjects tend to take into account the opinions of others. In particular, we find an interaction effect between social awareness and the display of user preferences. In particular, our findings suggest that people who tend to listen more to other users, tend to perceive better the explanations that include information about their own profile. On the other hand, when user preferences are not displayed, the perception of explanation quality remains pretty much the same, despite the extent of users' social awareness. At this respect, we believe that users with greater abilities to take into account the opinion of others might appreciate the chance to see the alignment of their own preferences with the opinions of others, in an effortless manner, since the metric of importance to the user for each aspect was placed next to the metric of other users' opinions for the same aspect.

On the other hand, even though the mean transparency is slightly higher for the "yes" condition of display user preferences variable, the differences are not significant. This is somewhat surprising, as we expected, in particular, that users would perceive as much more transparent a system that showed them which of their information the RS was using to infer their "user profile". However, it is difficult to make a strong statement about this based on the present study, mainly because the system did not display the real subjects' preferences (preferences were fixed as indicated in the cover story). Second, more skeptic users may think that the system is hiding additional information about the user's profile that might be used to generate recommendations, so showing only frequencies of user aspects' mentions may not be enough to satisfy their curiosity. Similar concerns apply to the confidence variable, where no significant differences between conditions were found either.

Additionally, results showed no significant differences in the perception of effectiveness, across conditions. However, the results suggest that the more rational a user is when making decisions, the more useful the proposed explanations will be in achieving their goal (choosing a hotel). In particular, we observe a tendency to further penalize visual explanations in terms of perceived effectiveness when users are less rational in making decisions. This could

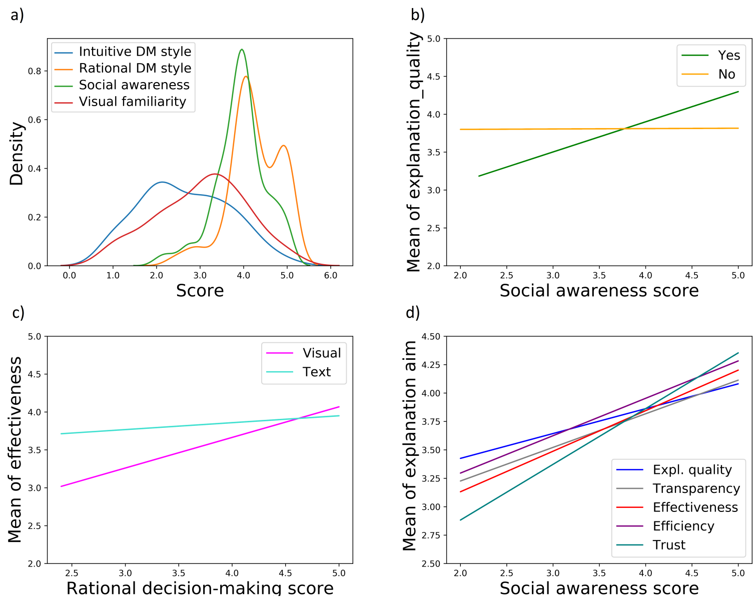

Figure 2: a) Kernel density estimate of user characteristics scores: rational and intuitive decision making styles, social awareness and visual familiarity. b) Interaction plot for explanation quality (fitted means of individual scores) between display of user preferences and social awareness. c) Interaction plot for effectiveness (fitted means of individual scores) between display style and rational decision making style. c) Effect of social awareness on all explanation aims (fitted means of individual scores). All scores within the range [1,5].

be explained by the fact that more rational subjects prefer to use as much information as possible to make an informed decision, while less rational subjects not only tend to need less information, but may find the visual content of the explanation too overwhelming or difficult to process compared to the condition of the text, so that the processing effort does not compensate for the amount of information required. The same pattern was also observed for the efficiency variable.

Finally, our results indicate that social awareness seems to play a meaningful role in the overall perception of both explanations and the overall RS, since we found significant main effects of it on all evaluated variables. Here, people with a higher disposition to listen and take into account others' opinions, tend to like the proposed explanations more, and to perceive the overall RS as more transparent, effective and trustworthy than people with lower social awareness.

7 CONCLUSION

In this paper, we have presented the design of argumentative explanations based on reviews, in display styles that involve both visual and text elements, as well as information about the user preferences. We also addressed the role that individual differences regarding decision making styles, social awareness and visual familiarity play in such perception. Although we found no main differences in perception between the regarded display styles, nor the presence or absence of user preferences in explanations, we found that, when taking into account user characteristics, i.e. social awareness, rational or intuitive decision making style, we are able to do detect differences in explanations' perception between users.

As a follow up study, and given the subtle differences on the effect of explanations when no individual differences are regarded, we will test the effect of every proposed design in a within subjects design. We believe that if users have the opportunity to contrast the possible advantages that each design provides, they could more distinctly indicate their true preferences, which would help us better understand in which cases each display has a more positive reception.

An important limitation of our study is the fact that user's preferred aspects were fixed and participants were instructed to pretend that those aspects were the ones that mattered most to them, aiming to give a practical work around to the cold-start problem in the user study design. However, we acknowledge that this might interfere with the real perceived benefit of providing the user preferences as part of the explanations, reason why we plan to provide a mechanism that allows participants to read explanations that fit to their real preferences, e.g. providing a preliminary view of a limited number of user profiles, representative of clusters of users with similar preferences extracted from dataset, so the participant could pick the user profile that fits best to her/his own interests.

REFERENCES

[1] Konstantin Bauman, Bing Liu, and Alexander Tuzhilin. 2017. Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 717–725.

[2] Chong Chen, Min Zhang, Yiqun Liu, and Shaoping Ma. 2018. Neural attentional rating regression with review-level explanations. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web. International World Wide Web Conferences Steering Committee*. 1583–1592.

[3] Felipe Costa, Sixun Ouyang, Peter Dolog, and Aonghus Lawlor. 2018. Automatic Generation of Natural Language Explanations. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. 57:1–57:2.

[4] Ruihai Dong, Michael P. O Mahony, and Barry Smyth. 2014. Further Experiments in Opinionated Product Recommendation. In *Case Based Reasoning Research and Development*. Springer International Publishing, 110–124.

[5] Collaborative for Academic Social and Emotional Learning. 2013. 2013 CASEL guide: Effective social and emotional learning programs - Preschool and elementary school edition. (2013).

[6] Katherine Hamilton, Shin-I Shih, and Susan Mohammed. 2016. The Development and Validation of the Rational and Intuitive Decision Styles Scale. *Journal of Personality Assessment* 98, 5 (2016), 523–535.

[7] 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*. ACM, 241–250.

[8] Diana C. Hernandez-Bocanegra, Tim Donkers, and Jürgen Ziegler. 2020a. Effects of Argumentative Explanation Types on the Perception of Review-Based Recommendations. In *Adjunct Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP '20 Adjunct)*. https://doi.org/10.1145/3386392.3399302

[9] Charles A. Hill and Marguerite Helmers. 1986. *Communication and persuasion: Central and peripheral routes to attitude change*. Springer-Verlag, New York.

[10] John R Kirby, Phillip J Moore, and Neville J Schofield. 1988. Verbal and visual learning styles. *Contemporary Educational Psychology* 12, 2 (1988), 169–184.

[11] Bart P. Knijnenburg, Martijn C. Willemsen, Zeno Gantner, Hakan Soncu, and Chris Newell. 2012. Explaining the User Experience of Recommender Systems. In *User Modeling and User-Adapted Interaction*. 441–504.

[12] Pigi Kouki, James Schaffer, Jay Pujara, John O'Donovan, and Lise Getoor. 2019. Personalized Explanations for Hybrid Recommender Systems. In *Proceedings of 24th International Conference on Intelligent User Interfaces (IUI 19)*. ACM, 379–390.

[13] Chin-Yew Lin. 2004. Rouge: A package for automatic evaluation of summaries. *Text Summarization Branches Out* (2004).

[14] D. Harrison McKnight, Vivek Choudhury, and Charles Kacmar. 2002. Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. In *Information Systems Research*, Vol. 13.

[15] Khalil Ibrahim Muhammad, Aonghus Lawlor, and Barry Smyth. 2016. A Live-User Study of Opinionated Explanations for Recommender Systems. In *Intelligent User Interfaces (IUI 16)*, Vol. 2. 256–260.

[16] Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*. Association for Computational Linguistics, 311–318.

[17] Pearl Pu, Li Chen, and Rong Hu. 2011. A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems - RecSys 11*. 157–164.

[18] Wolfgang Schnotz. 2014. Integrated Model of Text and Picture Comprehension. In *The Cambridge Handbook of Multimedia Learning (2nd ed.)*. 72–103.

[19] Paul Thagard and Abninder Litt. 2000. Models of Scientific Explanation. *The Cambridge handbook of computational cognitive modeling* (2000).

[20] Nava Tintarev. 2007. Explanations of recommendations. *Proceedings of the 2007 ACM conference on Recommender systems, RecSys 07* (2007), 203–206.

[21] Nava Tintarev and Judith Masthoff. 2012. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction* 22 (2012), 399–439.

[22] Stephen E. Toulmin. 1958. The Uses of Argument. (1958).

[23] Jesse Vig, Shilad Sen, and John Riedl. 2009. Tagsplanations: explaining recommendations using tags. In *Proceedings of the 14th international conference on Intelligent User Interfaces*. ACM, 47–56.

[24] Henning Wachsmuth, Martin Trenkmann, Benno Stein, Gregor Engels, and Tsvetomira Palakarska. 2014. A review corpus for argumentation analysis. In *15th International Conference on Intelligent Text Processing and Computational Linguistics*. 115–127.

[25] Nan Wang, Hongning Wang, Yiling Jia, , and Yue Yin. 2018. Explainable Recommendation via Multi-Task Learning in Opinionated Text Data. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 18*. 165–174.

[26] Yao Wu and Martin Ester. 2015. Flame: A probabilistic model combining aspect based opinion mining and collaborative filtering. In *Eighth ACM International Conference on Web Search and Data Mining*. ACM, 153–162.

[27] Yongfeng Zhang, Guokun Lai, Min Zhang, Yi Zhang, Yiqun Liu, and Shaoping Ma. 2014. Explicit factor models for explainable recommendation based on phrase-level sentiment analysis. In *Proceedings of the 37th international ACM SIGIR conference on Research and development in information retrieval*. 83–92.