Argumentation-Based Explanations in Recommender Systems: Conceptual Framework and Empirical Results

Sidra Naveed
Duisburg-Essen University
Germany
sidra.naveed@uni-due.de

Tim Donkers
Duisburg-Essen University
Germany
tim.donkers@uni-due.de

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

ABSTRACT

Explaining automatically generated recommendations has shown to be an effective means for supporting the user's decision-making process and increasing system transparency. However, present methods mostly provide non-personalized explanations that are presented in an unstructured manner. We propose a framework based on Toulmin's model designed to generate explanations in an argumentative style by presenting supportive as well as critical information about recommended items and their features. Existing research suggests that argumentative explanations cannot be assumed as equally effective for everyone. People rather tend to either apply rational or intuitive decision-making styles that determine which kinds of information are preferably taken into account. In an experimental user study, we investigated the effectiveness of argumentative explanations while considering the moderating effect of these two different cognitive styles. The results indicate that argumentative explanations, as compared to baseline methods, lead to, among others, increased perceived explanation quality, information sufficiency and overall satisfaction with the system. However, this seems only to be true for intuitive thinkers who rely more on explanations in complex decision situations as compared to rational thinkers.

CCS CONCEPTS

•Information systems → Recommender systems; • Human-centered computing → User Studies; **•Information systems → Personalization**

KEYWORDS

H.5.2 Information Interfaces and Presentation: User Interfaces-evaluation/methodology, graphical user interfaces (GUI), user-centered.*

ACM Reference format:

Sidra Naveed, Tim Donkers, Jürgen Ziegler. 2018. Argumentation-Based Explanations in Recommender Systems: Conceptual Framework and Empirical Results. In *Proceedings of UMAP'18 Adjunct, July 8-11, 2018, Singapore, Singapore, 6 pages.*
DOI: https://doi.org/10.1145/3213586.3225240

1 INTRODUCTION

Most of the current approaches of system-generated explanations for recommendations are limited in terms of the level of information they provide and the inability to justify the recommended items to users. Such simple explanations might not be sufficient for users in making their decisions especially, if the product domain is complex and financial or personal risk is associated with the purchase decision.

Such simple explanations have the potential to be more informative and personalized by explaining the relevancy of recommended items to users' personal preferences [11].

However, current techniques of personalized explanation are also limited in terms of providing structured reasoning and justification of the recommended items to users, as these explanations are composed of positive statements only with the purpose to persuade users rather than helping them in their decision-making processes.

Therefore, in the current work we go beyond the traditional approaches of explaining recommendations by proposing a conceptual model that provides personalized explanations in an argumentative manner. The process of argumentation is defined as an incremental selection of positive and negative statements to support or contradict a conclusion which in case of recommender systems (RS) is the system-generated recommendations. The main idea behind the proposed argumentative explanation is to not only provide the rationale behind the recommendations but also explain the (un)suitability of recommended items. These recommendations are justified progressively by generating a series of positive and negative statements to help users in their decision-making processes.

Additionally, the thinking style adopted by individuals is generally considered dependent on the type and the level of information provided to them, thus changing the way these individuals make their decision [5].

Therefore, we hypothesized in general that different levels of explanations could affect the users' decisions about accepting the recommender system differently, for rational and intuitive decision-makers.

In the present paper, we provide several contributions. First, we propose a conceptual framework on how argumentative explanation can be imbedded in the process of the user's interaction with a RS. We also show how such explanations may be provided in the user interface. As a main contribution we report on the results of an empirical study that provides insights for the impact of levels of explanations on overall system acceptance as measured through factors such as perceived recommendation quality, information sufficiency, usefulness, users' satisfaction with the system, and the influence of these qualities on users' behavioral intentions. We selected baseline and personalized explanations to be compared with the proposed argumentative explanation approach because of the incremental nature of argumentation and the input sources exploited in our conceptual framework, which is the extension of the baseline and personalized explanation approaches. We further provide evidence that the acceptance of the system in the presence of different levels of explanations is likely to be moderated by user's decision-making styles i.e., rational and experiential decision style, albeit in a manner that may not be self-evident to RS designers.

2 RELATED WORK

To date, numerous approaches to explaining system generated recommendations have been proposed in RS. However, current explanation techniques mostly lacks in providing justification and explicit reasoning to the user to help them in their decision processes [12], as these techniques rely mostly on exploiting limited information sources and incorporate only positive statements to generate their explanations.

More recently, in limited research, some methods have been proposed to integrate argumentation technique in RS [2, 3], mostly with a focus to improve the recommendation quality as its inference abilities can generate recommendations and its structured reasoning in a systematic manner. This ability of argumentation resolves conflicts between user preferences and recommendations, thus providing suggestions accompanied by series of arguments.

In this context, an argumentation-based music recommender system was developed in [6], where Defeasible logic programming (DeLP) was used to model user preferences in terms of rules which were then used to generate recommendations. The reasoning based on these rules were then converted into human understandable form to provide explanations of the recommendations. The argumentation scheme has also been applied in social recommender system [9], where the argumentation model is used to provide justification of the recommendations to the active user, based on the preferences of his/her neighbors.

Additionally, some limited work exists that incorporated the argumentation in RS with main focus to enhance system transparency by improving the quality of explanation of recommendations. In this context, a framework to generate

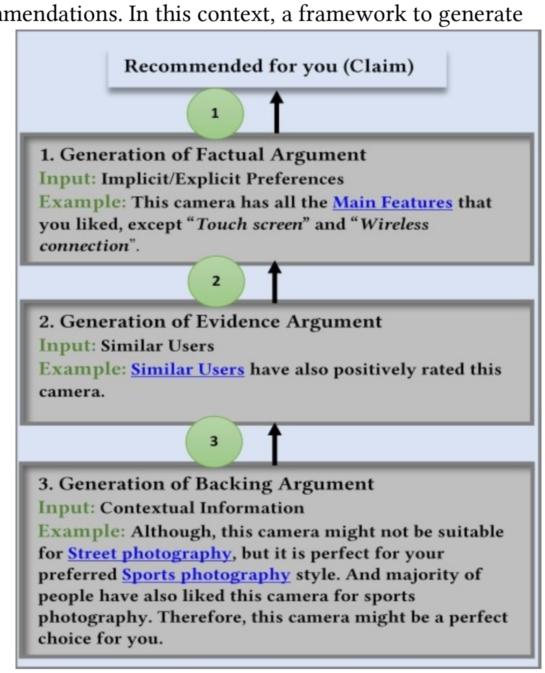

Figure 1: A framework for argumentation-based personalized explanation generation based on Toulmin's model of argumentation

personalized explanations based on arguments in a shopping RS have been proposed [11]. The system used the method of Multi-Criteria Decision Making (MCDM), to assess the quality of five pre-defined types of arguments to be included in the explanations. Text-based templates were used to generate explanations according to the kind of specified argument for the selected item.

A more similar approach was proposed to provide argumentative explanations for proactively delivered recommendations, in a gas station recommender scenario for automobiles [1]. The approach incorporated the ramping strategy to gradually provide various levels of explanations without distracting the driver. The Multi-Criteria Decision-Making Methods (MCDM) have been used to assess the quality of each argument that needs to be included in the explanation.

3 A FRAMEWORK FOR ARGUMENTATIION-BASED PERSONALIZED EXPLANATIONS AND UI DESIGN

In the current work, we propose a new conceptual framework

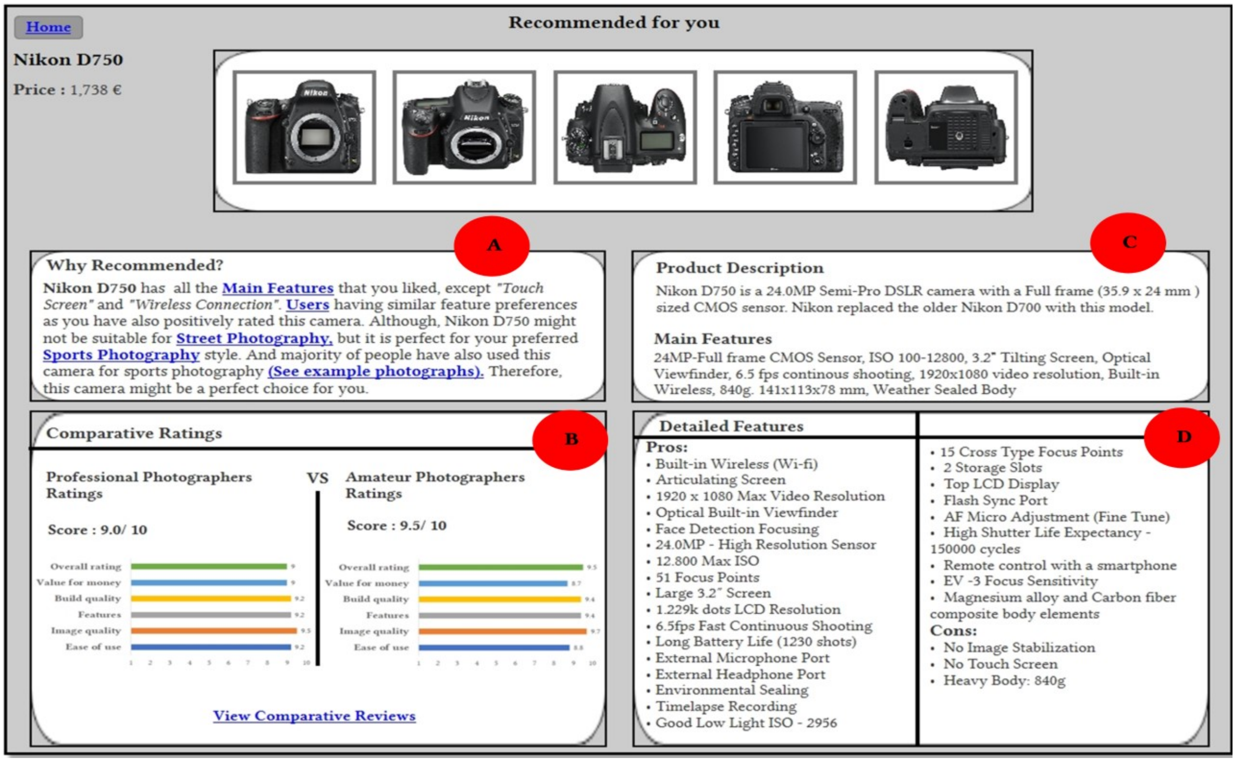

Figure 2: Explanation window where, (A) Argumentative-explanation area, (B) Comparative ratings and reviews, (C) Detailed product description, and (D) Detailed features.

with a main aim to integrate the argument-based reasoning into the RS to provide a qualitative perspective in decision-making. We used the basic structure of Toulmin's model of argumentation to construct our argumentative-explanations. The idea behind the Toulmin's model is to progressively establish the conclusion for the claim (which in case of RS are system generated recommendations) by generating series of positive or negative arguments that supports or contradicts the claim [7]. The sequence of these arguments is shown with numbers in green circle in Fig.1.

The arguments are generated by exploiting various information sources (i.e., user's implicit/explicit feedback, similar users, contextual information based on various photography styles and comparative opinions (in terms of ratings and review of social groups)). The first statement in the explanation is: (1) a *Factual argument* based on user's implicit/explicit preferences. This argument provides direct reasoning behind the recommendation, which is further supported by (2) an *Evidence argument* based on similar users and their preferences. The evidence justifies the relevance of the supportive data to the claim, which may be presented as a simple comment, or as a complex argument with additional sub-arguments. To further strengthen the evidence argument, (3) a *Backing argument* is provided by exploiting contextual information about various photography styles. Each highlighted text in the argument shows the argument premise that could be expanded by the user for detailed information about that argument.

Based on the proposed conceptual framework, we developed UI mockup design to show that how these argumentations maybe provided in the user interface. As our system did not yet generate real recommendations therefore, we selected five digital cameras to be presented as recommendations to users. Different web sources were used to extract information for digital cameras to present argumentations for the mockup design †. The workspace of our UI prototype consisted of two main windows i.e. Recommendations window and Explanation window. Due to space limitation, we have only presented explanation window and is shown in Fig.2., with different UI components marked with red alphabetical circles. The prototype had a limitation in terms of interactivity, allowing the users to only explore the items' details and explore the argument premises by clicking on the blue highlighted text in explanation area (A) and (B).

In Fig.2., (A)- Explanation area providing the series of positive and negative arguments about why the selected camera is suitable/unsuitable for the user. The sequence of arguments followed the steps of Toulmin's model, where each blue

† https://www.test.de/
http://cameradecision.com/
http://www.expertreviews.co.uk/

highlighted text opens a pop-up window, providing further explanation for that argument. These series of arguments were further supported by additional information in terms of comparative ratings and product reviews of professional and amateur photographers shown in region (B). (C) and (D) provides the detailed product description and features.

4 USER STUDY EVALUATION OF PROPOSED FRAMEWORK

We conducted a user study to examine the effectiveness of varying levels of explanations including our proposed argumentative explanation approach on user's acceptance of the system, where system acceptance is measured by various dimensions e.g., perceived recommendation accuracy and usefulness, explanation, information sufficiency, overall satisfaction, and purchase and use intentions. We also assumed that the user's decision-making styles i.e., rational or intuitive (moderating variable) would influence the effectiveness of these explanation approaches on overall system acceptance.

Put together, we formulated the following hypothesis:

H1. The overall *Acceptance of a RS* is dependent on the Levels of Explanations for recommended items with argumentative explanations being the most effective.

H2. The overall *System Acceptance* with different *Levels of Explanations* is moderated by the decision styles adopted by the user.

We further subdivided H2 into the following hypothesis:

H2.1. The overall *System Acceptance* for Argumentative explanations is higher than for Baseline and Personalized explanations for people with intuitive decision-making style.

H2.2. The overall *System Acceptance* is independent of Levels of Explanations for people with rational decision-making style.

4.1 Study Design

A total of 60 university students (45 female) participated in the study (Age: M=22.42, SD=4.48, range 18-35 years). A between-subjects experiment design was chosen. Each condition with n=20 differs in the level and type of explanations provided.

Baseline explanation is provided based on the top-rated cameras by the web users (e.g., *These cameras are recommended to you because these are the top DSLR cameras rated by users).* The personalized explanation is provided by exploiting user's past preferences and similar users' choices with only positive statements (e.g., *These cameras are recommended to you because it has all the main features that you liked and users having similar feature preferences as you have also positively rated these cameras).* The proposed approach as shown in Fig.1, provided argumentative explanation (including positive as well as negative statements) by exploiting not only user's past preferences and similar users' choices but also contextual information, photographers' reviews and ratings, thus extending the other two variants of explanations in an incremental manner.

Three UI design mockups for three variants of explanations (including our proposed argumentative-explanation as shown in Fig.2) were prepared for this user study to present recommended items and their explanations. While the mockups only allowed for simple interactive navigation and did not use actual recommender functionality, always showing the same set of recommendations, the design allowed to analyze the effects of different explanations independently from recommender performance.

4.2 Procedure

The user study procedure consisted of the following steps:

1. Participants provided their demographic information.

2. Participants recorded their responses on original English version of REI (40) [8] to calculate the moderating variables. The dimensions of rationality and experientiality are measured by aggregating the values of corresponding two-subscales i.e., Rational Ability-RA (10 items), Rational Engagement-RE (10 items), Experiential Ability-EA (10 items), and Experiential Engagement- EE (10 items)**,** where each item in the sub-scales are evaluated on a five-point Likert scale, from "completely false" (1) to "completely true" (5). Reverse coding is applied to all the negative statements, prior to calculating their scores.

3. Participants were given two decision problems of selecting a best matched digital camera from a list of five recommended cameras, based on the explicitly provided requirements in the two task scenarios. In both scenarios, the same list of cameras was shown to the participants. However, the two task scenarios differed in their complexity in terms of feature requirements.

4. Participants explore the system twice and find a camera after each interaction that satisfies each task scenario. Participants evaluated the overall system acceptance on items from the unifying evaluation framework, called ResQue (Recommender system's Quality of user experience) [13]. We used 12 constructs with 17 items from the ResQue and 5 additional items of our own for three constructs from the framework i.e., explanation, information sufficiency, and control, based on the relevance of these constructs with the scope and limitations of the study design. Each item was evaluated on a five-point Likert scale from "strongly disagree" (1) to "strongly agree" (5), where each construct is computed by averaging the scores of the composite items. The aggregated result from multiple constructs access the overall system acceptance from user's point of view.

4.3 Results

Table 1 provides a descriptive overview over the performance with respect to different dependent variables. The mean scores of the outcome variables are calculated based on the participants responses on five points Likert scale on ResQue questionnaire. The difference in three system conditions based on levels of explanations can also clearly be observed in Table 1.

4.3.1 Validating H1. To test the effect on overall system acceptance by the levels of explanation provided, we used MANOVA, with an alpha level of 0.05. The results revealed a statistically significant difference in the aggregated values of the dependent variables, phrased as *System Acceptance,* subject to *Level of Explanations* provided, $F(24, 92)=2.28$, $p=.003$*;* Wilk's $\Lambda= 0.394$, $\eta^2=0.372$. Multivariate effect sizes indicate that approximately 37% of the combined dependent variable is associated with the system condition factor.

To determine the effects of the individual dependent variables, we tested a series of follow-up ANOVAs. The results show that system condition has a statistically significant effect on Explanation (F(2, 57) = 15.79; p<.001; partial η^2=0.36). Post-Hoc tests indicate that *Baseline* explanations were perceived significantly worse than *Personalized* (p=.005) and *Argumentative* (p<.001). We also found a significant influence for Information Sufficiency (F(2, 57) = 4.18; p=.02; partial η^2=0.13) with Post-Hoc tests revealing that *Argumentative* yielded better results than *Baseline* (p=.015). In case of Use Intentions (F(2, 57) =6.39; p=.003; partial η^2=0.18), again, results indicate significantly better results for *Argumentative* as opposed to *Baseline* (p=.002). Finally, Overall Satisfaction (F(2, 57) =4.87; p=.01; partial η^2=0.15), was rated significantly higher in *Argumentative* than in *Baseline* (p=.009).

4.3.2 Validating H2.1. In order to validate H2.1, we utilized MANOVA to seek moderating effects of *Experientiality*. Multivariate tests indicate significant main effects on *System Acceptance* for *Level of Explanations*, F(18, 2)=41.73, p=.024; Wilk's Λ<0.001, multivariate η^2=0.997, as well as *Experientiality*, F(306, 31.07)=3.43, p<.001; Wilk's Λ<0.001, multivariate η^2= 0.968. However, more importantly in this context, it was found that the interaction between *Level of Explanations* and *Experientiality* was also statistically significant, F(126, 22) = 3.8, p<.001; Wilk's Λ<0.001, multivariate η^2= 0.929. This interaction effect indicates that experiential decision-making style moderates the relationship between *Level of Explanations* and overall *System Acceptance*.

To further determine the effect on the individual moderating effects, we conducted follow-up ANOVAs. To identify more precisely the nature of moderating effect, multiple comparisons using Post-Hoc tests were also made between means of three groups.

In terms of *Level of Explanations* individually, we, again, found significant effects for Explanation (F(2,9)=10.24; p=.005; partial η^2=0.7), Information Sufficiency (F(2,9)=8.69; p=.008; partial η^2=0.66), Use Intentions (F(2,9)=6.61; p=0.017; partial η^2=0.6), as well as Overall Satisfaction (F(2,9)=13.78; p=.002; partial η^2=0.75). Additionally, with the moderator included, we could now observe a significant effect on Perceived Ease of Use (F(2,9)=12.58; p=.002; partial η^2=0.74). The moderator *Experientiality* itself has a significant effect on Perceived Ease of Use (F(34,9)=3.73; p=.021; partial η^2=0.93), and Overall Satisfaction (F(34,9)=3.54; p=.025; partial η^2=0.93).

The interaction effect becomes significant for Explanation (F(14,9)=3.675, p=0.028; η^2=0.851). Post-Hoc tests reveal that, under the influence of the moderator, differences subject to *Level of Explanations* become even more prominent. *Personalized* (p=.004) and *Argumentative* (p<.001) were both rated better than *Baseline* while *Argumentative* was also preferred to *Personalized* (p=0.031). Furthermore, we found a statistically significant interaction effect for Information Sufficiency (F(14,9)=3.07, p=.048; η^2=0.827). In this case, Argumentative was perceived better than Baseline (p=.003). Another significant interaction effect could be found for Perceived Ease of Use (F(14,9)=3.08, p=.048; η^2=0.827) with significant Post-Hoc differences between *Personalized* and *Baseline* (p=.023). Finally, the interaction for Overall Satisfaction (F(14,9) =5.05; p=0.01; η^2=0.887) was also significant. Besides *Argumentative* (p=.001), *Personalized* explanations were also rated higher than *Baseline* (p=.013). Concluding, these results support H2.1.

4.3.3 Validating H2.2. To validate H2.2, we utilized MANOVA to seek a moderating effect of *Rationality*. Multivariate tests results indicate insignificant effects on *System Acceptance* for *Level of Explanations*, F(24,6) =1.0, p=0.572; Wilk's Λ=0.020, multivariate η^2=0.857, and *Rationality*, F(396,62.9)=1.06, p=0.407; Wilk's Λ<0.001, multivariate η^2=0.853. However, more importantly in this context, it was found that the interaction of varying levels of explanations and Rationality was also insignificant, F(132,33.2)=0.945, p=0.603; Wilk's Λ<0.001, multivariate η^2=0.687. This interaction effect indicates that Rational decision-making style does not moderate the relationship between *Level of Explanations* and *System Acceptance*, thus supporting the null hypothesis.

Table 1: Results of the outcome variables for the three groups who used Baseline, Personalized, and Argumentative versions of the system (Mean (SD))

Variables	Baseline	Personalized	Argumentative
Rec. Accuracy	3.87 (0.72)	**4.07** (0.54)	**4.07** (0.56)
Perc. Usefulness	3.55 (0.85)	3.5 (0.76)	**3.96** (0.48)
Confidence & Trust	3.48 (0.81)	3.5 (0.63)	**3.8** (0.62)
Interface Adequacy	3.36 (0.58)	3.53 (0.76)	**3.57** (0.56)
Explanation	2.87 (1.08)	3.75 (0.75)	**4.35** (0.58)
Inf. Sufficiency	3.5 (0.84)	3.85 (0.63)	**4.2** (0.79)
Perc. Ease of Use	3.9 (0.91)	**4.45** (0.68)	4.15 (1.08)
Control	3.3 (1.05)	3.67 (0.69)	**3.85** (0.48)
Diversity	3.35 (0.81)	3.7 (0.8)	**3.8** (0.95)
Use Intentions	2.55 (1.07)	3.13 (0.85)	**3.56** (0.75)
Purchase Intentions	3.25 (1.29)	3.35 (0.93)	**3.75** (1.01)
Overall Satisfaction	3.3 (1.12)	3.6 (0.82)	**4.15** (0.58)

The table shows that the argumentative condition outperformed the other two system conditions for most of the outcome variables, except the perceived ease of use where the personalized condition is better than the other two conditions.

5 DISCUSSION

The presented user study offers empirical evidence on how various levels of explanations in a complex risk-involved decision task affect perceived quality of RS and its recommendations for different decision makers. The main findings indicate that a user's attitude towards the system is greatly affected by the level of explanations provided. Concretely, concerning H1 we found that argumentative explanations performed better in terms of system acceptance. We assume that one critical reason for argumentative explanations being preferred is that they provide a well formulated rationale behind recommendations which the other two variants lack. Beyond that, the explanation gets an even more solid fundament by listing possible positive and negative consequences of a particular decision as shown by area (A) mentioned in the red circle, in Fig.2. Such polarized enumerative depictions can easily be grasped by people making complex decision domains more accessible and therefore increasing the system's trustworthiness [2].

Out of the 12 dependent variables that contribute to overall system acceptance, four came out to be statistically significant, i.e. Explanation Quality, Information Sufficiency, Use Intentions, and Overall Satisfaction. Our results are in accordance to existing research stating that argumentative explanations are designed to yield higher results in, for instance, explanation quality, information sufficiency, and overall satisfaction. However, receiving increased levels of Use Intention is surprising, especially when considering that interaction was merely performed on a design mockup. One possible reason might be that argumentative explanations offer direct textual assurance that a recommended camera is adequate for the intended use case, e.g. *"Nikon D750 is perfect for your sport photography needs"*.

Non-significant differences regarding the remaining eight dependent variables maybe also due to the system being only a mockup and due to a somewhat reduced statistical power as well as relatively small sample size.

Beyond the influence of systematically varying the level of explanations, H2 was concerning with the degree to which one relying on an explanation is also moderated by individual decision-making strategies. People with intuitive or experiential thinking styles showed more dependency on explanations to make advantageous decisions in a risk-involved situation. By contrast, those with rational or logical thinking styles seem to make their decisions independent of whether or not they are provided with explanations for their choices. We take this as evidence that the benefit of argumentative system feedback in the form of explanations is, in fact, moderated by participants' cognitive functions. Users with experiential thinking style deserve more guidance by the system, e.g. in the form of explanations, when they are about to make risk-involved decisions. These findings validate hypotheses H2.1 and H2.2 and are in line with research in psychology and cognitive sciences [4, 10].

6 CONCLUSION AND FUTURE WORK

The current work has made several contributions. First, we proposed a conceptual framework that extended the conventional explanation approaches to a more argumentative manner following the basic structure of Toulmin's model of argumentation. We presented an initial UI design mockup to show that how these argumentative explanations may be provided to users in complex decision situations. As a main contribution, we conducted a user study to investigate the impact of varying levels of explanations including the proposed argumentative explanation, on overall system acceptance. The study results validated our conceptual framework, where the argumentative explanations outperformed the other two variants of the explanations in terms of better system acceptance by users. Our result findings further showed that the acceptance of system in the presence of varying levels of explanations is moderated by decision styles of the user i.e., rational and experiential. The results indicated that the users with experiential thinking style showed more dependency on the levels of explanations provided to them as compared to the rational decision makers.

The future work will focus more on methods to develop a real recommender system based on our proposed framework. Furthermore, we will focus more on developing and demonstrating the explanation algorithms, i.e., methods for selecting, structuring and evaluating the arguments to be included in an explanation. Additionally, we tend to explore methods to present these textual argumentations in a more interactive and visual manner without overwhelming the user with information overload.

REFERENCES

[1] Bader, R., Woerndl, W., Karitnig, A. and Leitner, G. 2011. Designing an explanation interface for proactive recommendations in automotive scenarios. *International Conference on User Modeling, Adaptiation, and personalization (UMAP)* (2011), 92–104.

[2] Bedi, P. and Vashisht, P. 2014. Empowering recommender systems using trust and argumentation. *Information Sciences.* 279, (2014), 569–586. DOI:https://doi.org/10.1016/j.ins.2014.04.012.

[3] Bedi, P. and Vashisth, P. 2011. Interest based recommendations with argumentation. *Journal of Artificial Intelligence.* 4, 2 (2011), 119–142. DOI:https://doi.org/10.3923/jai.2011.119.142.

[4] Brand, M., Fujiwara, E., Borsutzky, S., Kalbe, E., Kessler, J. and Markowitsch, H.J. 2005. Decision-Making Deficits of Korsakoff Patients in a New Gambling Task With Explicit Rules: Associations With Executive Functions. *Neuropsychology.* 19, 3 (2005), 267–277. DOI:https://doi.org/10.1037/0894-4105.19.3.267.

[5] Brand, M., Laier, C., Pawlikowski, M. and Markowitsch, H.J. 2009. Decision making with and without feedback: The role of intelligence, strategies, executive functions, and cognitive styles. *Journal of Clinical and Experimental Neuropsychology.* 31, 8 (2009), 984–998. DOI:https://doi.org/10.1080/13803390902776860.

[6] Briguez, C.E., Budán, M.C.D., Deagustini, C.A.D., Maguitman, A.G., Capobianco, M. and Simari, G.R. 2012. Towards an argument-based music recommender system. *COMMA. Frontiers in Artificial Intelligence and Applications.* 245, 1 (2012), 83–90. DOI:https://doi.org/10.3233/978-1-61499-111-3-83.

[7] Van Eemeren, F., Garssen, B., Krabbe, E.W., Snoeck Henkemans, a F., Verheij, B. and Wagemans, J.M. 2014. Toulmin's Model of Argumentation. *Handbook of Argumentation Theory.* Springer, Dordrecht. 203–256.

[8] Epstein, S. 1973. The Self-Concept Revisited: Or a Theory of a Theory. *American Psychologist.* 28, 5 (1973), 404–416. DOI:https://doi.org/http://dx.doi.org/10.1037/h0034679.

[9] Heras, S., Navarro, M., Botti, V. and Julain, V. 2009. Applying dialogue games to manage recommendation in social networks. *International Workshop on Argumentation in Multi-Agent Systems* (2009), 256–272.

[10] Kahneman, D. 2011. *Thinking, Fast and Slow.* Macmillan.

[11] Lamche, B., Adigüzel, U. and Wörndl, W. 2014. Interactive explanations in mobile shopping recommender systems. *In Joint Workshop on Interfaces and Human Decision Making in Recommender Systems* (2014), 14–21.

[12] Nunes, I. and Jannach, D. 2017. A systematic review and taxonomy of explanations in decision support and recommender systems. *User Modeling and User-Adapted Interaction.* 27, 3–5 (2017), 393–444. DOI:https://doi.org/10.1007/s11257-017-9195-0.

[13] Pu, P. and Chen, L. 2011. A user-centric evaluation framework of recommender systems. *Fifth ACM conference on Recommender systems* (2011), 157–164. DOI:http://doi.org/10.1145/2043932.2043962