Multi-Criteria Rating-Based Preference Elicitation in Health Recommender Systems

Helma Torkamaan
University of Duisburg-Essen
Duisburg, Germany
h.torkamaan@acm.org

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

ABSTRACT

A multi-criteria rating looks for important dimensions to more extensively capture an individual's opinion about a recommended item. Health Recommender Systems (HRS) is considered to be an emerging domain of recommender systems. In HRS, criteria for a multi-criteria preference elicitation of a recommendation have not yet been fully investigated to the best of our knowledge. In this paper, we investigate the criteria for the rating of a health promotion recommendation using an online survey. Drawing on both the relevant literature and the users' responses, we came up with a list of 33 criteria that users are considering when they rate a health promotion recommendation. However, these criteria are not equally important to users. We discuss which of these criteria are more important in the users' opinions. In short, our results show that users consistently consider effectiveness, emotional gain, and giving a good feeling as the most important criteria. Using the criteria derived from the literature, we came up with a model for the importance of the criteria which has three dimensions: effect, effort, and context. This study is the first step toward enhancing our understanding of HRS and the rating of a health promotion recommendation.

CCS CONCEPTS

• **Information systems** → *Information systems applications*; *Recommender systems*; *Information retrieval*; • **Human-centered computing** → *Human computer interaction (HCI)*; • **Applied computing** → *Health care information systems*; *Health informatics*;

KEYWORDS

Recommender systems, Health recommender systems; Multi-criteria rating; Health promotion

ACM Reference Format:
Helma Torkamaan and Jürgen Ziegler. 2018. Multi-Criteria Rating-Based Preference Elicitation in Health Recommender Systems. In *Proceedings of the Third International Workshop on Health Recommender Systems co-located with Twelfth ACM Conference on Recommender Systems (HealthRecSys'18), Vancouver, BC, Canada, October 6, 2018* , 6 pages.

1 INTRODUCTION

Recommender systems (RS) have used various methods and algorithms that utilize implicit and explicit user feedback in order to build up a user profile. While such systems may obtain implicit user feedback via modeling user behavior, explicit user feedback mainly relies on the ratings that a user gives to either the recommended or consumed items. With a user-oriented perspective toward RS, one may consider that in real-life, users may use various dimensions to describe their attitude toward a product or recommendation. For instance, a user may say: *I like the durability of the product, but it is not aesthetically beautiful*, or she might say: *I like how the actor is playing in this movie, but I hate the ending.* Therefore, it is important to consider how a rating feedback for a recommended item should be, how it should be interpreted, and how it is reflected in the algorithm and recommendation process.

Researchers have looked into one-dimensional or single-criterion rating and later also into multi-criteria rating. In single-criterion rating, user feedback is captured via a single value from a bounded set of natural numbers. The user feedback here represents either an overall rating of a recommended item or an alternative question that one may ask when the rating task is presented to the users. Even though single-criterion rating has been commonly used, over the years researchers [2, 5, 8, 10] have pointed out its limitations in capturing the users' subjective opinions and drawn our attention to multi-criteria rating in RS. Unlike single-criterion rating, multi-criteria rating tries to find important dimensions for capturing a broader spectrum of an individual's opinion toward the recommended item, and accordingly increases the quality of the recommendations using this extra information about the user's preferences [1, 6]. Adomavicius and Kwon [1] maintain that multi-criteria RS were developed not only to give more accuracy and present the complexity of users' preferences, but also to fulfill the challenges associated with multi-objective recommendation strategies or multiple performance criteria of an RS.

Health recommender system (HRS), as an emerging domain of RS, represents three classes of solutions: I. the application of RS in the health (medical) domain, II. the application of RS in the health promotion and well-being domain, and III. RS that exploit health information in order to give recommendations in other domains. The field of HRS, as an interdisciplinary domain, deals with various challenges [11] and some of these challenges, such as persuasion, may fall beyond typical RS challenges. One may argue that HRS is both a multi-stakeholder and multi-objective domain. Accordingly, it is important for the HRS community to consider multi-criteria rating and find out what dimensions are essential in rating if one explicitly rates a recommendation.

While one can obtain the criteria for multi-criteria rating from both previous research and the collected datasets in domains such as movies or tourism, in the health domain accessing similar datasets is not easily possible and such criteria, therefore, have not been

HealthRecSys'18, October 6, 2018, Vancouver, BC, Canada

fully investigated. As a result, criteria for this particular domain remain unclear. Even for those available datasets with a multi-criteria rating, it is up to researchers to investigate how important each dimension is and construct a consistent group of criteria. Such criteria enable the scientific community to come up with suitable algorithms for multi-criteria RS. The present study reports on an exploratory attempt to find out the dimensions or criteria that are influential in the rating of a recommended item in the HRS domain, focusing on the second category of HRS. This paper also reports on which criteria users consider the most important and ultimately, proposes a model for user satisfaction with a recommendation. The result of this attempt can be used to construct a primary set of criteria for health-specific domains and would help the RS community to deepen their understanding of user satisfaction and rating.

2 METHOD

An exploratory study was conducted using an online survey in order to investigate the criteria. In addition, we sought to investigate the importance of the obtained dimensions or criteria for the overall rating in the users' opinions. Participants answered the questions regarding various recommendation scenarios related to a smartphone health application. For this study, we considered four aspects: I. the study should be health-domain-dependent II. various scenarios should be considered when asking for a user rating, III. familiar scenarios with mobile app mockups should be presented to the participants, so they can easily rate them, and IV. study design should prevent bias related to the criteria and their importance by ordering questions appropriately and ensuring the sequence of the given criteria changes randomly.

The online study contains four sections with the following order: subject knowledge, scenario, importance ranking, and demographics. While the demographic section asks for standard questions about gender, education, age, country of origin, and native language, subject knowledge section asks participants if they have ever rated a product or wrote a review on a product online. The subject knowledge section also determines to what extent a participant has made decisions on consumption of a recommended item based on its rating.

The scenario section included three scenarios. A recommendation (social engagement, physical activity, and positive psychology) follows each of the scenarios with the aim of mental health promotion and reduction of the negative effects of stress in daily life. The survey shows a mockup containing the recommendation on the smartphone, and then asks the participants to follow the recommendation. They are then asked if they did follow the recommendation and subsequently, they rate the given recommendation overall. After the rating, the participants are asked to write down the criteria and the reasons that they have considered for giving the overall rating in an open-ended question. In another open-ended question, the survey asks participants if the ratings are to be used to provide personalized recommendations for them, which factors they would consider while rating a health promotion recommendation. We designed these open-ended questions, particularly to capture criteria that users consider for giving a rating to a recommendation.

The section then continues by showing the previously shown scenarios and recommendations again. This time, the participants are asked to give separate ratings for the recommendations in the following criteria: location, fitting personal preference, suitable time, meeting their goals, being interesting, enjoyable, being effective, suitable time consumption, emotional aspect, easy to do, suitable cost, and socially acceptable. We obtained this list of criteria from the literature. For instance, using both Fogg's behavior model for persuasive design [4] and Health Belief Model [9], we listed cost, time-consumption, effort, difficulty, social deviance and emotional gain and effectiveness criteria. We included location and time factors as an extension to the Fogg's behavior model, related to the signal as trigger considering that if a signal as trigger arrives at the wrong time or location, it can easily be ignored by the users.

Other factors have their origin in the literature as well. For example, easy to do refers to a task not being too difficult and not being too easy for a user inspired from both self-efficacy theory [3] and goal-setting theory [7]. Meeting the goals is also inspired from goal-setting theory [7]. Fitting personal preferences comes from the user's explicit preference from RS domain. We also added being interesting and enjoyable as additional factors.

The next section of the study, the ranking section, asks participants to rate a set of given criteria (randomly ordered) based on the importance a criterion has in the participant's overall rating of a health promotion recommendation. We included 19 items (listed in the left column of Table 2) by extending the previously mentioned criteria. For example, we had two items of *emotional gain or positivity* and *good feeling*. In addition, to re-check for the most important criteria in the participants' opinion, participants ranked the eight top criteria in order of importance in another question. The survey was online for the duration of 5 days.

3 RESULTS

Participants' description. A total of 74 (36 females, 30 males, and 8 prefer not to say) participants completed our study. The age range of the participants is between 20-54 years old with 32.5% between 20-30 and 56.8% between 30-40 years old. Among the participants, 87.8% had university degree education. The participants' origin is from 14 different nationalities. While 79.7% of participants have previously rated a product or service online, only 66.2% have written reviews. Among the participants, 75.6% consider online rating important and very important in their decision making on the consumption or purchasing of a product or service and another 23.0% consider the rating moderately important. Finally, 72.9% of the participants very often and always check online ratings when they are looking for a product or services and another 17.6% sometimes check the ratings.

Descriptive results. Qualitative analysis of 296 open-ended answers in total gives us the insight for criteria that were not included in our 19-item criteria. Interestingly, we were able to easily code the answers into the 19-item criteria described earlier in the method section so that only a few instances of a new criterion were left out. In addition to the 19-item criteria, the participants mentioned the explanation of the recommendation and its origin (n=13) and the recommendation being new or not new (n=26) as criteria for rating a health promotion recommendation. The participants also mentioned the following criteria in less than five instances: empowering, trust, transparency, suitable for my current mood,

Table 1: The most important criteria in the participants' view in the rating of a health promotion recommendation, Q1- ranking task (Rank one is the best, the lower the value the better the rank), Q2- importance rating 1-10, a higher value describes more importance.

Criteria	Mean -Q1	SD-Q1	Mean-Q2	SD-Q2
Emotional gain or positivity	3.29	2.18	8.57	1.822
Giving a good feeling	3.41	2.36	8.47	1.854
Effectiveness	3.9	2.05	8.57	1.912
Interesting	4.00	2.42	8.19	2.01
Fitting personal preferences	4.28	2.02	8.21	2.10
Fulfilling my goals of using the app	4.33	2.16	8.35	2.15
At a suitable time	4.72	2.20	8.33	1.88
I enjoy the recommendation	4.93	2.15	8.53	1.85
Would follow it	4.57	1.63	7.57	2.18
Not time-consuming	4.57	2.47	7.49	1.96

encouraging, not-human-like (getting advice from a machine), understandable, easy to read, needing it, intrusiveness, funny, and relevance for me. The participants in general mentioned these criteria from the 19-item criteria more than others: effectiveness of a recommendation (n=97), the emotional gain (n=66), fitting one's preferences (n=47), liking (n=43), at a right time (n=32), in a right location (n=25), and easy to do (n=19). In total, all criteria including repeating similar items create a list of 33-items.

Although in the open-ended questions we observed a difference in the frequency of the mentioned criteria between different scenarios, in all three scenarios as well as in the general question, *effectiveness* of the recommendation is the most frequently mentioned criterion. *Emotional gain* for the two scenarios of positive psychology and social engagement was the second most frequently mentioned criterion. At the same time for the physical activity scenario, the second most frequently mentioned criterion was *at a suitable time*. *Fitting one's preferences* and *liking* as well as *would do the recommendation* criteria were mentioned frequently in the physical activity and social engagement scenarios, however, these criteria were not among five top frequently mentioned criteria for the positive psychology scenario. Instead, participants mentioned *not difficult*, *meeting my goals*, and *not time-consuming* criteria for the positivity psychology scenario.

The overall rating and criteria-based rating of the recommendations, in all scenarios together, show positive correlations ($p < .01$) with the following criteria: effectiveness ($r_s = .68$), emotional gain ($r_s = .61$), fitting personal preferences ($r_s = .65$), being interesting ($r_s = .62$), and enjoyable ($r_s = .57$). Other criteria are also significantly related to the overall rating ($.2 < r_s < .44$).

Two questions considered the importance rating of the criteria: the ranking question (Q1) and the importance rating (Q2). The ranking[1] result shows that the participants as a whole relatively consider three criteria of *giving a good feeling* ($M = 3.41, SD = 2.36$), *emotional gain or positivity* ($M = 3.29, SD = 2.36$), and *effectiveness* ($M = 3.90, SD = 2.05$) to be the most important ones. The analysis of the importance rating question, Q1, with the range of importance score being between *(not important at all)1-10 (very important)* only confirms top three important criteria, Table 1. Among other criteria, *fitting personal preferences*, *being interesting* as well as *fulfilling my goals of using the app* all had mean values among the top eight mean values in both Q1 and Q2, however, there is a slight difference for the last two sorted items in Q1 and the results of Q2. In Q2, *the recommendation at a suitable time* and *I enjoy the recommendation* have better mean values while in Q1, *would follow the recommendation* and *the recommendation is not time-consuming* have better mean values.

An exploratory factor analysis (i.e. maximum likelihood using Promax rotation) on 19 criteria from Q2 was conducted. Three factors were extracted that accounts for 58% of the variance overall. The loading of each of the criteria on these factors is presented in Table 2. Based on the intuitive interpretation of the analyzed item, we summarize the factors as the recommendation effect (F1), the recommendation effort (F2), and the recommendation context (F3). These three factors show three important aspects of the rating of a health promotion recommendation. While *effect* represents the influence of the recommendation on the users including the recommendation effectiveness and its emotional consequences, the recommendation *effort* represents the challenges that a user may need to overcome in order to follow the recommendation. This aspect may also deal with the motivation and goal setting. Finally the third factor, the *context* represents the physical context of a user such as their location and time of the recommendation. The internal consistency of the factors was assessed using Cronbach's α for the factors F1, F2, and F3 and we obtained the following values respectively .947, .811, and .839 that is acceptable.

Table 2: The level of importance a criterion has in general for giving an overall rating for a health promotion recommendation. The criteria are in the left column. The loadings of the criteria are in the right columns. High loading (>0.30) is in bold.

Variable	F1	F2	F3
Liking	**0.42**	**0.39**	0.02
Doing	**0.52**	0.08	0.12
Suitable time	0.20	0.07	**0.68**
Not interrupt	-0.08	0.07	**0.69**
Not cost money	-0.01	**0.67**	-0.28
Not much effort	0.07	**0.40**	0.30
Socially acceptable	-0.01	0.29	**0.37**
Good feeling	**0.89**	0.04	-0.03
Fulfill goals	**0.66**	-0.10	0.24
Fits preferences	**0.70**	0.07	0.17
Emotional gain	**0.92**	0.00	-0.01
Not difficult	-0.12	**0.90**	0.13
Effective	**0.99**	-0.17	-0.04
Not time-consuming	-0.25	0.30	**0.60**
Right location	0.19	-0.26	**0.91**
Gender role	-0.04	**0.49**	-0.03
Interesting	**0.74**	0.19	-0.06
Enjoying	**0.90**	-0.03	0.04
Easy to do	0.05	**0.91**	-0.06

[1] Q1: the most important being 1 to 8

4 DISCUSSION

The results of our study in the demographic and subject-knowledge sections reveal that our participants were familiar with RS and ratings as users, and in fact, they were making some of their purchasing and searching decisions by relying heavily on such ratings. The variety of nationalities and ages in our study represents a wider range in comparison to the college-student population. These sections leads to the assumption that the participants were familiar with both giving a rating and the concept of a recommendation based on the ratings. Consequently, we may assume that our participants were able to reflect on the rating criteria in the questions.

The open-ended questions suggest criteria in addition to our 19-item criteria extended from the literature. These results confirm our design of 19-item criteria and add to it. Some criteria mentioned by participants such as trust or transparency are very important particularly in HRS despite being mentioned only in a few instances by the participants. Lack of trust in an HRS domain can lead to disastrous consequences such as harm to individual or leak of sensitive data. Furthermore, users also have a higher expectation of privacy, trust, and transparency for such systems in comparison to other RS domains. However, one may also consider that a criterion such as *the recommendation explanation* can have an effect on these criteria. It is therefore important for the HRS community to consider the explanation aspect as a criterion alongside designing trust-aware and privacy-aware systems.

A criterion such as *being new* represents two contrasting groups of participants. Some participants appreciated the familiarity of the recommendation and emphasized that they would like to be reminded of the recommendation that they had prior positive experience with. These participants mentioned that they would follow the recommendation, and they like the recommendation and its effectiveness due to their prior positive experiences. In contrast, some participants gave a low rating for the recommendation despite expecting a positive effect from the recommendations. These participants mentioned that they do not need the recommendation, because they already do the recommendation on a daily basis. Such a difference only highlights the importance of personalization of a recommendation. For example, imagine in specific domains of HRS that a behavior change design requires user engagement in a repetitive behavior, particularly for such a design, the researchers should be aware of the user preferences and their assumptions about giving a rating. Alternatively, one could also consider specific engagement approaches or motivations such as emotional gain to increase the user satisfaction and compliance alongside instructing the users for giving a rating feedback.

From all the mentioned criteria, *effectiveness* seems to be the most important criteria. However for HRS in health promotion or behavior change, instantaneous effectiveness may not be observable for the users. In fact, the effectiveness depends on the users' goals of using the HRS application and its design or their prior experiences. A user may need to follow a program to the end to feel a change or improvement, or they could even feel worse right after initiating a behavior change process, for example, in a smoke cessation behavior change. Therefore, a challenge would be to show the users an effect. Differentiating between long-term and short-term effect and planning for instantaneous effect could be a possible solution. In particular, users do not always consider a final goal as the effectiveness and therefore a design for an instantaneous emotional gain (hope, pleasure, reward, etc.) may give the users the effect they would like to receive.

Emotional gain or positivity and *good feeling* were also frequently mentioned as important criteria. While these two criteria in our opinion represent the same aspect, i.e. emotional gain, other criteria also may contribute to it. Imagine criteria such as liking, interesting, or enjoying; they all contribute to the users' final emotional gain or loss feeling. Even motivational criteria such as a task being too difficult or too easy can indirectly lead to the feeling of pleasure. Criteria such as social acceptance can lead to the feeling of shame or fear. Interruption, suitable time or location can also lead to a higher cognitive load or even sometimes turn into a hassle. Accordingly, it is expected that emotional gain criterion is very important in the rating of a recommendation and one should consider that other criteria at the end may influence this criterion.

We intentionally designed different scenarios to capture various criteria. Initially, we thought that these scenarios may lead to various ratings of criteria. Nevertheless, except for some minor differences, the overall result did not suggest a major difference in the ratings of the most important criteria being effectiveness and emotional gain, user preferences, interesting, and fulfilling the goals of using the application. For instance, criteria related to F3: the context, namely, *at a suitable time*, *not interrupt*, *not time-consuming*, and *right location* were expected to show more importance for the users. One could explain this in two ways. It is plausible that the context criteria are more important on the decision-making of the user on following the recommendation versus ignoring it rather than on its overall rating since these criteria may play a role as a signal as the trigger in Fogg's behavior model [4]. In this case, the context criteria perhaps influence more on the users' overall satisfaction of the application and HRS rather than the individual rating of a recommendation. One could therefore assume that the user's rating of a criterion reflects only on the criteria in effect factor, and accordingly, criteria in both effort and context factors would lead to the behavior of following or ignoring a recommendation. It is also possible that the limitations of an online survey instead of a real-world testing prevent the participants' reflection on the context criteria. We intend to follow up this preliminary finding with the application developed for this purpose. We would test in a real-life situation and checking the rating feedback, would possibly confirm which of the mentioned possibilities are the case.

In either possibility, the HRS designers could rely on using more than just explicit ratings to capture the users' feedback. For instance, one can easily capture the context criteria using user behavior tracking or use just in time interruption techniques to minimize interruptions. Alternatively, one could also ask users if they would follow the recommendations and only capture user feedback for the consumed items. The effectiveness of a recommendation may also be captured using indirect methods. For example in a physical activity recommendation, one can easily capture the changes using the activity trackers or for stress reduction, one can assess before and after level of stress. Overall, the feedback mechanism and the use of rating in HRS depends on the targets of an HRS as well as the study design. Imagine that you want to use a specific behavior change model or a set of intervention that may not focus on the

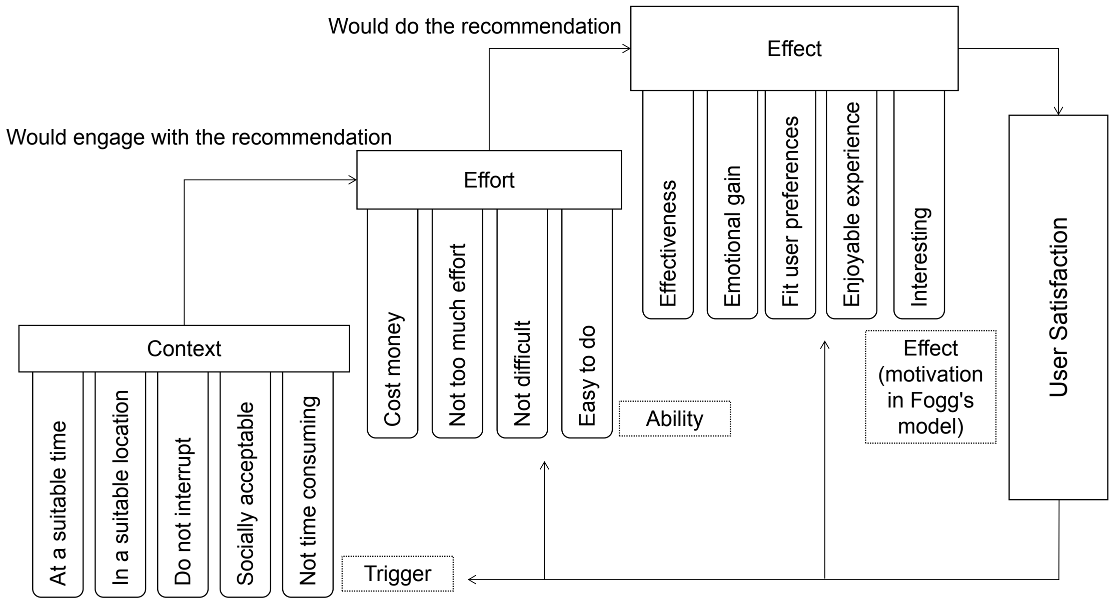

Figure 1: Effect, Effort, Context (EEC) model describing the user satisfaction of a health recommendation

user's rating at all. When the design of HRS allows, a design using a multi-criteria rating feedback may allow a better personalization and a higher effectiveness and user satisfaction as a result.

4.1 A basic model for user satisfaction with a recommendation

We propose a basic model for user satisfaction with a health promotion recommendation. Considering both the three extracted factors for importance and the importance rating of the criteria, and taking inspiration from Fogg's behavior model that encompasses motivation, ability, and trigger [4], one could observe three dimensions that are essential in user satisfaction with a recommendation. Extracted factors of F1: effect, F2: effort, and F3: context respectively would represent reward or goal, motivation, and signal, Figure 1.

Context. A user **would engage with a recommendation** if it is suggested in a convenient situation or physical context for the user. A convenient situation means at the right time, in the right location, for the right duration, with no inconvenient interruption, or with the recommendation being suitable considering the presence of others near the user. If the context is not convenient for the recommendation, then regardless of its effect, it could be simply ignored.

Effort. A user **would engage with a recommendation** if it falls within her capacity or ability to do or cope with the task. If the effort condition is not met, then the recommendation may be ignored or would possibly yield unsatisfactory results. One should, however, consider that these two dimensions of effort and context are preconditions that could determine if the user will follow a recommendation to the end.

Effect. A user accordingly **would be satisfied by a recommendation** if the recommendation has an effect on her, including either emotional gain, pleasure, or effect, or fulfilling her explicit or implicit goals and motivations. These goals could be a final target or a behavior change, or be as simple as a reward or pleasurable experience. All three components together are required for a recommendation to satisfy a user. First, the user should be able to do the recommendation, or it should be probable that the user can follow the recommendation. Then she should engage with the recommendation to the end, and finally, the recommendation should bring the user the *effect.*

In this model, any of the context criteria can disrupt fulfillment of the context dimension. For the effort dimension, a combination of the criteria may or may not disrupt its fulfillment, and only the capacity of an individual user would describe this combination. Any of the criteria in the effect dimension may complete its fulfillment and lead to a higher user satisfaction. Effect not only influences user satisfaction, but it can also shape the users' preferences. For example, if a user has a good prior experience with a recommendation, she might decide to follow the recommendation with an even greater effort. On the contrary, an unsatisfactory effect could possibly lead the user to ignore the recommendation right away. As tagged in figure 1, the context dimension is parallel with the trigger factor in Fogg's behavior model. The effort dimension is parallel with the ability factor, and the effect dimension is parallel with the motivation factor in Fogg's model. Nevertheless, Fogg's model describes a behavior model for a persuasive design and is not a model of user satisfaction from a recommendation.

5 CONCLUSION AND FURTHER WORK

Multi-criteria rating provides an opportunity to gain a fuller picture of an individual's opinion. This paper explored the criteria for the domain of health recommender systems and provided 33 criteria that users are considering when rating a health promotion recommendation. Finding out that these criteria are not equally important for the users, this study determined the criteria that are more important for users than others. Using a 19-item criteria list, this work presented a three-dimensional model that represents the aspects that these criteria refer to. Our work clearly has some limitations. The online survey is limited in its ability to capture real-life user responses and reactions. In addition, criteria such as item explanation and novelty were not included in the factor analysis since they were captured in the same survey. Despite this, we believe that our work could be a starting point for developing multi-criteria recommendation algorithms for HRS. The extracted list of criteria gives an insight into the criteria that are important in the users' ratings of a health promotion recommendation.

REFERENCES

[1] Gediminas Adomavicius and YoungOk Kwon. 2015. *Multi-Criteria Recommender Systems.* Springer US, Boston, MA, 847–880. DOI:http://dx.doi.org/10.1007/978-1-4899-7637-6_25

[2] Gediminas Adomavicius, Ramesh Sankaranarayanan, Shahana Sen, and Alexander Tuzhilin. 2005. Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach. *ACM Trans. Inf. Syst.* 23, 1 (Jan. 2005), 103–145. DOI:http://dx.doi.org/10.1145/1055709.1055714

[3] Albert Bandura. 1977. Self-efficacy: toward a unifying theory of behavioral change. *Psychological review* 84, 2 (1977), 191.

[4] BJ Fogg. 2009. A Behavior Model for Persuasive Design. In *Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09).* ACM, New York, NY, USA, Article 40, 7 pages. DOI:http://dx.doi.org/10.1145/1541948.1541999

[5] Dietmar Jannach, Fatih Gedikli, Zeynep Karakaya, and Oliver Juwig. 2012. Recommending Hotels based on Multi-Dimensional Customer Ratings. In *Information and Communication Technologies in Tourism 2012.* Springer, Vienna, 320–331. DOI:http://dx.doi.org/10.1007/978-3-7091-1142-0_28

[6] Dietmar Jannach, Zeynep Karakaya, and Fatih Gedikli. 2012. Accuracy Improvements for Multi-criteria Recommender Systems. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC '12).* ACM, New York, NY, USA, 674–689. DOI:http://dx.doi.org/10.1145/2229012.2229065

[7] Edwin A Locke and Gary P Latham. 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist* 57, 9 (2002), 705.

[8] Nikos Manouselis and Constantina Costopoulou. 2007. Analysis and Classification of Multi-Criteria Recommender Systems. *World Wide Web* 4, 10 (2007), 415–441. DOI:http://dx.doi.org/10.1007/s11280-007-0019-8

[9] Irwin M Rosenstock. 1974. Historical origins of the health belief model. *Health education monographs* 2, 4 (1974), 328–335.

[10] Fernando Sanchez-Vilas, Jasur Ismoilov, FabÃŋn P. Lousame, Eduardo Sanchez, and Manuel Lama. 2011. Applying Multicriteria Algorithms to Restaurant Recommendation. In *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01 (WI-IAT '11).* IEEE Computer Society, Washington, DC, USA, 87–91. DOI:http://dx.doi.org/10.1109/WI-IAT.2011.124

[11] Hanna Schäfer, Santiago Hors-Fraile, Raghav Pavan Karumur, André Calero Valdez, Alan Said, Helma Torkamaan, Tom Ulmer, and Christoph Trattner. 2017. Towards health (aware) recommender systems. In *Proceedings of the 2017 international conference on digital health.* ACM, 157–161.