User Groups and Different Levels of Control in Recommender Systems

Christine Mendez[1], Vlatko Lukarov[1], Christoph Greven[1], André Calero Valdez[2], Felix Dietze[2], Ulrik Schroeder[1], Martina Ziefle[2]

[1]Learning Technologies Research Group, RWTH Aachen University
Ahornstraße 55, 52074 Aachen, Germany
[2]Human-Computer Interaction Center, RWTH Aachen University
Campus-Boulevard 57, 52074 Aachen, Germany
christine.mendez@rwth-aachen.de, lukarov@cil.rwth-aachen.de, greven@cs.rwth-aachen.de, calero-valdez@comm.rwth-aachen.de, dietze@comm.rwth-aachen.de, schroeder@cs.rwth-aachen.de, ziefle@comm.rwth-aachen.de

Abstract. The aspect of control in recommender systems has already been extensively researched in the past. Quite a number of studies performed by various researchers reported that an increase in control had a positive effect for example on user satisfaction with a system, or recommendation accuracy. Recent studies investigated whether this positive effect of control applies to all users, or finer distinctions have to be made between different user groups, which in turn require different levels of control. Those studies identified several characteristics, along which users could be divided into groups: expertise in recommender systems, domain knowledge, trusting propensity, persistence. They reported different needs of control for different user groups. However, the effect of those characteristics has not been systematically examined with regard to all three recommendation phases introduced earlier by Pu and Zhang, namely *initial preference elicitation, preference refinement, result display*. This paper suggests, that for different levels of expertise and trust, different levels of control are necessary during preference elicitation, whereas persistence does not play a prevalent role in this phase. Further assumptions are made for preference refinement and result display. In addition to the three phases, *context, type of information required* and *visualization of control methods* are identified as factors influencing the request of users for control.

Keywords: Recommender Systems; User Groups; Controllability; User Satisfaction

1 Introduction

The research focus in the area of recommender systems has shifted from recommendation algorithms, to the users and their needs [16]. Various studies covering the aspect of control in recommender systems reported a positive effect of control on user satisfaction on various systems (e.g. [11], [2], [1], [10]). In general, user involvement and

adfa, p. 1, 2011.

interactive interfaces do have immediate impact and effect on the user experience. So far, to our knowledge, only a few studies examined the general assumption that more control automatically leads to better user experience. A more fine-grained distinction between user-groups revealed, that different levels of control are requested by different users ([12], [7], [9], [8]). In those studies, researchers distinguished users by varying levels of the following characteristics: expertise in recommender systems, domain knowledge, trusting propensity, and persistence. In fact, they found that different user groups responded differently to an increase in control. This paper provides an overview over past results by classifying them according to the phases of the recommendation process. For each phase we will discuss suggestions for further empirical research and formulate hypotheses based on previous research.

2 Controllability in Recommender Systems

One way to introduce controllability in recommender systems is to actively involve the user in the recommendation process (workflow). The user can shape and improve his profile within the system; change and adapt the way recommendations are presented to him; provide additional context information which the system cannot implicitly collect in order to receive more suitable suggestions. These are just a small set of user activities in different stages/phases of the recommendation process which increase controllability. In the following section we will provide detailed overview of different interaction techniques and activities, and how were they used in different phases of the recommendation process.

2.1 Ways to Control the Recommendation Process

Looking at existing work that deals with controllability and user interaction with recommender systems, it is clear that there are different ways to allow the user to influence the recommendation process. Some notable examples are creating a user profile by "giving binary or multi-scale scores" to items, tagging items, weighting of item attributes, critiquing recommendations, The list of intervention possibilities throughout the recommendation process goes on and on [13].

In order to facilitate the comparison of different controlling capabilities, we will classify them with respect to key activities in the recommendation process or phases identified by Pu and Zhang: (i.) (Initial) Preference Elicitation. In the initial phase the system gains initial knowledge about the user and establishes accurate interest profile in order to recommend items that match the user's taste. [13] (ii.) Preference Refinement. In this phase, the system refines the user's preferences after the initial phase of recommendation. In this phase, the user has the opportunity to alter (update) his preferences in order to receive more appropriate recommendations from the system. Activities like adding supplementary ratings to sample items, or evaluating the recommendations themselves via critiquing can help the recommender system from going in the wrong direction with the recommendations. (iii.) Result Display Strategies. In this phase, the generated recommendations are presented to the user [13]. In the following sections we will present exemplary studies, which implemented control mechanisms for each of the three phases. Fig. 01 provides an overview of these exemplary studies. It illustrates how those examples support the different recommendation phases and summarizes their findings on how those control mechanisms affect the user.

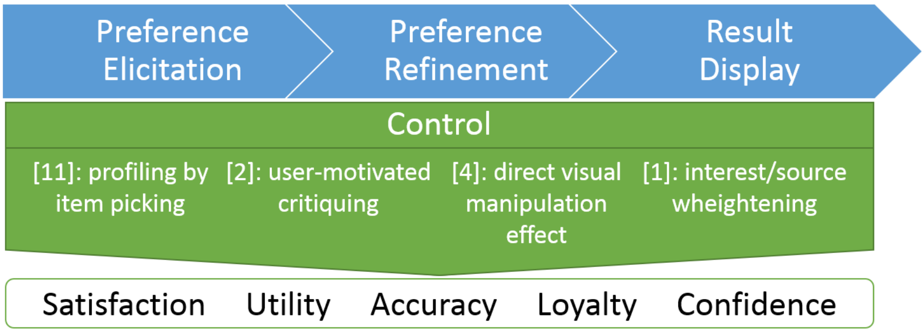

Fig. 1. Three recommendation phases with examples of control mechanisms and correlations to several positive aspects.

During Initial Preference Elicitation

In their study, McNee et al. [11] compare varying amounts of user control in the sign up process for the online movie recommender platform “MovieLens”. New users are asked to rate 10-15 movies they have seen, in order to create a user model and receive better recommendations. They found that the users may not only be involved by rating movies suggested by the system (system-controlled approach), but also the users themselves can suggest and rate movies they had seen in the past (user-controlled approach).

They discovered that the user models created with the user-controlled interface were more accurate than the others and that the users were more satisfied with their recommendations than those in the other focus group. Furthermore, participants in the user-controlled group tended to return to the system and showed a more active use after the sign-up process [11]. McNee et al. call them more “loyal” to the system and attribute this to their personal involvement in the initial preference elicitation. Although the sign-up process in this condition took the users almost twice as long, they did not perceive it as being lengthy. This, McNee et al. argue is also due to their involvement with the system [11].

Loepp et al. [10] present another way to exploit user interaction in the initial preference elicitation process. They specifically stress that their approach helps in overcoming the cold-start problem and can be applied for a user who does not have or want a permanent profile [10]. In this particular case, the user picks different sets of items, instead of rating individual items. The increase of user-provided information leads to positive user feedback w.r.t. fit, novelty of the results, control, effort and adaption [10] in the recommendation process, when compared to the manual search or automated recommendations with no user interaction. Lastly, they derived from their observations that their approach is especially well suited for users who do not have formed a concrete search goal yet [10].

During Preference Refinement

One example for an increase in control during preference refinement is discussed by Chen and Pu [2]. They examine different kinds of critiquing-based recommender systems. Generally speaking, these systems allow the user to improve or narrow down their results by critiquing the recommendations they received earlier. Chen and Pu compare dynamic critiquing (as a form of system-proposed critiquing) with user-motivated critiquing. In this comparison they analyze both objective (decision accuracy, task completion time, interaction effort) and subjective (perceived cognitive effort, decision confidence, trusting intentions) criteria [2]. They observe that user-motivated critiquing required less effort and resulted in a higher decision accuracy. Also this system was perceived as less demanding in terms of critiquing effort by their participants. The perceived critiquing effort was also correlated with the users' confidence in the results. The higher the user's confidence, the lower their perceived effort [2].

Gretarsson et al. [4] developed an interactive visualization that allows the user to explore their Facebook friends' interests and thus pick recommendations or results of interest: SmallWorlds. In their design, they graphically depict the user surrounded by alternating layers containing both their friends and their friends' items of interest, which are unknown to the user so far. Initially, the distance between a user and a friend is defined by the amount of common interests. The more items they have in common so far, the closer a friend (and their respective interests). Dragging a friend or an item nearer to the user increases their weight and highlights related items that might match the user's taste.

Gretarsson et al. [4] found that this mapping for the process of weighting or selection is very intuitive and easy to handle. Furthermore, their results indicate that people could easily find common tastes. The same holds for popular and interesting items which were easy to identify. Their participants also stated that using SmallWorlds was easier than browsing a text-based interface for common interests and new recommendations. The parallel comparison of their friends was perceived as an exceptional advantage.

Upon Result Display

Bostandjiev et al. [1] developed a system called TasteWeights. Although they also rely on Facebook data to create recommendations, their interface and interaction possibilities differ from the SmallWorld concept. Their surface is divided up into three columns. The first column contains a user profile with their (Facebook) items listed. The second column consists of three parts, representing three different recommendation engines which use the profile data to come up with recommendations. The recommendations are displayed in the third column. Each column allows the user to control the weight of an item or a recommendation technique. The impact on the other columns are immediately displayed on the graphical interface. For their study, Bostandjiev et al. compared the user satisfaction and recommendation accuracy when they were allowed to manipulate a set of various interaction possibilities in the system.

They found that the possibility of interacting with all three columns (which also includes an immediate (or direct) observation of the impact on the recommendations) lead to the most accurate recommendations and was perceived useful in the recommendation process. The full interaction condition also outperformed hybrid recommendation approaches using the same engines but offering no interaction to the user [1].

Ekstrand et al. [3] also let users control the algorithm with which their recommendations are created. While TasteWeights allowed the user to weight the contribution of a recommender engine to the result set, Ekstrand et al. let the user choose one algorithm which creates the whole result set.

They found, that not even a third of the users they recorded, used the opportunity to switch between different recommendation algorithms to get different result sets. Those who did, tended to use the system rather actively in general [3]. One can argue that such behavior indicates (but does not prove) that there might be a correlation between active users and the need of control. Examining only the users who switched between recommenders, it became clear that they tend to experiment with the obtained results early during a session, but settle for one algorithm rather quickly (no further changes in later sessions/during that session). For those users, user satisfaction and recommendation accuracy is high [3].

2.2 User Reactions to Levels of Control

As the examples show, there are several ways of giving more control to the user during all phases of the recommendation process. The general results show that an increased level of control is well-appreciated by the users. An increase in control lead for example to increased recommendation accuracy ([11], [2], [1]), fit, novelty [10], and thus user satisfaction ([1]) and a better feeling of control and adaption ([11], [10]) in various studies. Moreover, Chen and Pu could report that more control in critiquing-based recommenders cost their users less effort [2]. A similar result was obtained by Parra and Brusilovski [12]: The majority of their participants stated that using their visual control mechanisms facilitated their task compared to a baseline interface with no control options.

Harper et al. [5] also reported positive feedback to increased control levels in their experiment. Users report that their “adjusted lists better represent their preferences” and “subjects responded positively to a survey question if they would use a [control] feature” like the one used in the study [5]. This feedback is particularly interesting since users also reported usability issues with the given system. This suggests that the benefit of the control options outweighed inconveniences during the use of the tools.

Such positive results lead Hijikata et al. [6] to examine whether the mere feeling of being in control is enough to obtain higher user satisfaction during the use of a recommender system, without actually intervening in the process. In their study they compared two user groups: The first group using actual control mechanisms, while the other users were presented with ‘placebo controls’ which had no effect on the recommendations. They found that actual control correlated positively with user satisfaction, whereas the fake controls showed no such effect. [6]. This is an important result in so far as it supports the claim, that the user's participation in the recommendation process is a valid contribution to its results. This underlines the importance of using tools for including the user in the recommendation process.

Starting from those general results, a few attempts were made to distinguish between several user groups and their individual needs of control. Researchers started posing the question whether control is actually desirable for all kinds of users, or if there are certain characteristics which lead to different user reactions. Table 1 gives an overview of user categorizations from past research work. Different user groups are distinguished from the uniform crowd according to certain user characteristics. Those groups are then separately confronted with different levels of control. By doing so, the results reflect a more diverse population with users who have different prerequisites and needs. Concrete results of past work will be presented in the next section.

Table 1. Influence of control increase on different user groups, based on [8] and [4].

user category	user type	cause
expertise (with RS)	novice	☺
	professional	☺
domain knowledge	beginner	☹
	expert	☺
trustfulness	sceptics	☺
	trustful	☺
satisficer/maximizer	satisficer	☹
	maximizer	☺

3 Distinguishing User Groups and their Specific Needs

When developing an interactive system in general, the two most important questions one needs to answer are: Who are the users, and how are they going to use the system? In the following sections we try to categorize the different user groups and identify their specific needs within the context of controllability of recommender systems. We considered different factors, such as, information type, context, goal, interface or results visualization, and others, when distinguishing the different user groups and addressing their specific needs.

3.1 Previous Work

Jameson and Schwarzkopf [7] question the assumption that more control leads to more user satisfaction. In their study they found that both automatic and manual updating of recommendations on a hotlist had supporters amongst their participants. They identify the following factors that must be taken into account when giving control to users [7]:

1. The nature of the application and of the adaptation involved
2. Individual differences among users in terms of preferences, experience and ways of approaching the tasks in question
3. Contextual factors like speed of internet connection
4. Random situational factors like the nature of the information retrieved during the process.

Parra et al. [12] introduced a hybrid recommender system called SetFusion, which gives control to the user during both preference refinement and result visualization. It allows to adjust the weight of each of the different recommendation approaches that contribute to generating results. Furthermore, the results are visualized in an interactive Venn diagram which allows further inspection of an item and filtering the results according to certain criteria (for further details, see [12]).

They examined the effect of those control opportunities on user engagement and user experience with the system. In addition to the work of Jameson and Schwarzkopf, they considered concrete user characteristics: (i) User expertise (domain), (ii.) User experience with the underlying system, (iii.) trusting propensity, (iv) user experience with recommender systems in general [12].

They found that "past expertise of different kinds appears to be an important factor", influencing the user experience [12]. Their users showed more engaged use of the recommender system in the examined use case. Moreover, they perceived a higher diversity in the results [12]. This can be seen as contributing to a positive user experience. Parra and Brusilovski also found that a higher trusting propensity leads to increased trust in the visualization tools provided. What is interesting is that they found further characteristics like gender and native language to correlate with certain behavior [12], which is to be analyzed further in the next section.

In their online study, Knijnenburg et al. [9] rely on the TasteWeight system for social recommendations. They allow for different levels of control (none, item-control, friend-control) and inspectability (display the whole graph or just a list of results during the inspection phase). Their results show that both control conditions lead to higher system understandability with a notably lower inspection time. This ultimately leads to a higher perceived quality of the recommendations. Furthermore, they discovered a positive effect of trusting propensity on user satisfaction and a negative effect of expertise on user control. But although experts tend to feel less in control, they had higher ratings for perceived recommendation quality and satisfaction with the system [9]. In [8] they compare user satisfaction with different interaction methods: (i.) no interaction: list of items by popularity, (ii.) sort: sorting a given list by an attribute (from a list of possibilities), (iii.) explicit: weighting attributes, (iv.) implicit: having weights assigned according to their browsing behavior, (v.) hybrid: combination of (iii.) and (iv.). Each of those interaction techniques corresponds to a concept of decision making [8]. They then tested which interaction method was preferred by users with different levels of certain characteristics: domain knowledge, trusting propensity and persistence.

Indeed, Knijnenburg et al. could confirm their prediction that there was no overall best system, but that users preferred different systems according to their user group. For example, novice users had higher results for perceived control with the first system (where actually they did not have any opportunity of manipulation), whereas experts experienced the least control in this condition. A surprising finding was that the level of perceived control between the other three systems did not differ significantly for experts. Furthermore, as Knijnenburg et al. expected, experts showed the highest user interface satisfaction with the hybrid system, which was the least satisfying for novices [8].

The results for trusting propensity were different from what had been expected beforehand. Different controlling conditions provoked no significant difference for the measured understandability of the system. There was however a positive correlation between trusting propensity and perceived system effectiveness for method (i.) no interaction, (iii.) explicit, (iv.) implicit. Regarding the user interface satisfaction, trusting propensity correlated positively with the explicit and implicit interface. Persistence, i. e the span between maximizers who strive for the optimal result w.r.t all aspects of their research and satisficers who are content with the first result which matches an acceptance criterion, showed no effect on understandability, perceived effectiveness, control, or user interface satisfaction for any of the conditions [8].

3.2 Further Possible Factors to Distinguish Appropriate Amount of Control

After reviewing previous work on different user-groups and their control preferences, we would like to point out further possible areas of research which were not covered by previous studies. Also further factors were identified which, in certain circumstances, might have an influence on the amount of control desired by the user.

Recommendation phase: In the second chapter, different recommendation methods were introduced and assigned to the three key activities - initial preference elicitation, preference refinement, and result displaying - defined by Pu et al. [13]. In order to make a more fine-grained distinction between user needs, we propose that those phases in the recommendation process themselves function as factors which influence the adequate amount of control for a certain user group. We already presented that different recommendation methods offer different opportunities for the user to intervene in the recommendation process and thus control its results. We propose that those ways of interaction are not equally well suited for different user groups, as will be examined further in this chapter. Finally, we will take a closer look at previous findings from different papers and compare and relate their outcomes. Based on this theoretical analysis, further guidelines are proposed for the implementation of control in recommender systems which still might need further confirmation by evaluations.

During Initial Preference Elicitation

In the studies presented in this section, no further differentiation between users was made in order to better understand their needs of control in the first phase of preference elicitation. We took the user characteristics that were examined in the other recommendation phases (preference refinement and result display) and reflected upon their possible interaction with control methods and other relevant factors in the first phase. This reflection leads to the following outcomes.

In preference elicitation the type and nature of the required information has a significant impact on creating the user profile, or creating the item's characteristics within the recommender system. This information can be classified in three categories: personal information, domain-specific information, and context information.

Personal information: If the initial preference elicitation does not require any prerequisite knowledge, novices, as well as, experts are capable of intervening in the recommendation process by creating a profile and providing personal information. This activity requires low level of cognitive effort. Hence, one could argue that both user groups would appreciate a high level of control. This could be achieved by similar methods to those used by McNee et al. [11] who allowed participants to come up with their own items (in terms of interests, characteristics, etc.) instead of having them rate given options during creation of their profile. This would be an easy and understandable way for users with different levels of domain knowledge to exert control and refine their results. In order to facilitate the task, one could also provide optional suggestions, as in the hybrid interface of [11].

Users who are rather skeptical might also appreciate elaborate control methods that allow them to specify what information to share and what to keep to themselves, as opposed to more trusting users.

Domain-specific information: This type of information also elicits and contains personal preferences, but at the same time this kind of information is more specific to their domain of knowledge and operation. In their study, McNee et al. [11] required domain-specific information, when they asked to either rate given films, or to come up with films one has already seen and then rate those. As their results show, users who were given more control in their study tended to be more loyal to the system in a later stage of use. Nevertheless, this focus group had the most participants that gave up on the sign up process. Further examinations of this correlation could provide new knowledge. A possible finding could be (a) that users who were given more control yielded better results and invested more energy in the recommendation process. Both factors that could explain their observed loyalty. One could also argue that (b) given the high rate of people who gave up in the sign up process, only those users who are highly interested in the domain (in this case 'music') were willing to invest more effort into initial preference elicitation. Their loyalty could thus be attributed to their interest in the domain and not to the sign up- and recommendation process.

Based on (b) we would argue that during preference elicitation if domain-specific information is required, experts value an increase in control more than novices.. This suggests that during preference elicitation, it would make sense to confront the user with a "slimmed down interface" which requires a minimal level of interaction, but gives the user the freedom to provide more information, putting in extra effort. The user should always have the opportunity to give up some personalization for more comfort and ease of use. The user thus could cut right to the chase, or take time to configure their profile. If the system in use is designed for a user to come back, it could hint that a later refinement is always possible and/or ask the user if they wanted to make more configurations when they sign in the next time. During the next phase, novices should be given the opportunity to make more elaborate adjustments as their expertise in the domain and the system itself increases (see also next section).

With such two-fold interface, the user may choose their preferred way of preference elicitation. Instead of letting users control content, they are able to control how an entire step in the recommendation process is designed. We are thus proposing a different kind of control method which focuses less on results, but more on the user and their satisfaction with the system. Again, this solution might also be appealing to persons who are less trustful, since they do not have to fill in information, they consider inappropriate.

In this context, we would like to recall the following guideline: The effort a user must invest in the system must be lower than the necessary effort to solve the task manually by himself [12], [11]. In the light of the previous argumentation it is important to differentiate between user groups when it comes to the term effort or "cognitive effort". Cognitive effort is composed differently depending on the user's domain knowledge. A novice already has to put effort in order to cope with the information and causalities inherent to a problem, also referred to as its cognitive load, e.g by Sweller [15]. Experts, on the other hand, know their way in the domain and have more capacities to invest into system specific control structures. This is why a novice can only be expected to handle a lower amount of control than an expert to achieve their respective goal. This is why, especially on their first encounter with the system, they should be in the position to decide on the complexity they think they can handle. Users could also differ in the goal they want to achieve. To use again the example of a music recommender, the goal to "find new music that I like" could mean different things to different user groups. The criteria (domain specific) that have to be met in order to yield a satisfying result are probably more complex and diverse for an expert than for a novice. This aligns with Chen and Pu [13] who suggest in their design guidelines to optimize preference elicitation "to favor a small effort in the initial sign-up process" [13]. Their idea to implement an "incremental preference elicitation method" (see preference refinement phase) is picked up and expanded further by not only allowing them to apply the same intervention methods over and over, but also by providing them with an increasing set of tools for refinement. In our opinion this supports the idea of letting users control part of the recommendation process and their (item specific) control opportunities in order to achieve their respective goals.

Regarding persistence, we would argue that the amount of control available for the user does not play such an important role in the initial recommendation phase. As maximizers and satisficers mainly differ in their interaction with a given set of results, one can conclude that their preferences of control might differ more in a later phase.

Context: Apart from the user preferences, Ekstrand et al. [3] mention "context" as a factor influencing the choice of an appropriate recommendation algorithm. In their study they determine the user's context by relating to their behavior. This means that they exploit indirect user information in order to improve their choice of algorithm. We would consider context more generally as an important factor in the recommendation process. Since during one "recommendation session", the context does not change, or if it does the change is gradual and in most cases user-inspired, we consider it appropriate to define context during preference elicitation. All subsequent iterations of refinement and filtering can be based on the initial definition.

The definition of a context of use is another opportunity for a user to be in control. Contrary to [3], we would suggest allowing users to characterize their context the same way as they are allowed to characterize themselves e.g. in their profile for user modelling. As the user knows best in what context they use the recommender, it seems natural to involve them in the definition of the context created by the recommender system. Especially for distrusting users, it might be a more agreeable option to control the amount of information they share about their context of use (e. g on a mobile device, for private use, at a certain location, etc.) instead of feeling supervised, or tracked by a system. As with personalization, one could consider offering the user to contribute information or to explicitly allow services like GPS (as seen with mobile apps), tracking user behavior or not to use context information at all. Again, the amount of control could be variable and should be adjustable at any time.

During Preference Refinement

The second phase is the one that was investigated the most with respect to user groups in previous studies. Characteristics like expertise (in both domain and recommender systems), persistence or trusting propensity and their relation with control in recommender systems were already analyzed by Parra and Brusilovski [12] for example. As described earlier, they analyzed those characteristics using SetFusion, which implements several control methods using visual representation and feedback. Knijnenburg et al. [9] investigated in a similar direction: They also examined the level of users' trust, persistence, and expertise using the TasteWeights recommender system. This system also allows the user to interact on a graphical interface and gives visual feedback.

Interface: Although those studies already gave insight into the relation between user characteristics and control, there are several possibilities where further research can be done. First of all, as hinted above, the recommender systems examined in [12] and [9] both give control to the user in the form of graphical tools like sliders and diagrams. Critiquing-based recommenders, as described by Chen and Pu [2] follow a different approach which is more based on textual representation of information. It might be worth investigating how given user characteristics like trusting propensity, expertise (recommenders / domain), and persistence influence the amount of control desired in such a system where control is implemented in a different form.

We suggest that, overall, users feel less in control in a more text-based interface, than in a graphical interface like [12] and [9]. We base this statement on the fact that in the graphical approach not only the controls are represented graphically, but also their impact is mostly immediately visible on screen (see levels of inspectability in [9]). This parallel display of controls and effect differs notably from critiquing-based recommenders. Those use a rather iterative approach: A list of results can be filtered (more or less flexible) by critiquing single elements and a new list is generated which is then displayed. The relation between a concrete critique of an attribute is harder to track in the newly generated list because it puts cognitive load on the user, and on the other hand, changes cannot be reverted as easily as in [12] and [9].

Nevertheless, we would predict that more control, as exemplified by Chen and Pu in their "example critiquing interface" [2] would be appreciated by both novices and experts. Their example critiquing interface allows freely combining critiquing units to customized compound critiques. As they explain themselves, similarity-based critiquing ("give me similar items"), quality-based critiquing ("items like this but cheaper") and quantity-based critiquing ("items like this but $100 cheaper") can all be modelled through their interface [2]. This allows a user to be vague and unspecific in their critiques, depending on their goals and previous knowledge. In this situation, novices have the choice of starting with critiquing units or they can freely combine them. Compared to predefined compound critiques, this avoids the fear of changing undesired values, by picking a compound critique that does not fit well. For experts the freedom of choice w.r.t attributes and values gives them the opportunity to leverage their domain knowledge in order to find the most fitting result in few steps.

The process of filtering on the basis of a more personally tailored set of results, which is possible with increased control in critiquing-based recommender systems should be highly appreciated by maximizers, who have a high level of persistence. With the opportunity to fine-tune given results in any possible way, their striving for "the best" possible outcome should be best supported with a high amount of control as in the example critiquing interface [2]. User with a lower level of persistence might appreciate this opportunity as well, but will probably be equally satisfied when presented only with less control like in the dynamic- (system-proposed-) critiquing interface.

Goal / Context of Use: As during the process of preference elicitation, it could also be beneficial to give a sort of "meta control" to the user in the phase of preference refinement: instead of letting the user control attributes of the recommended items, it might be helpful to give the user a choice on how preference refinement is implemented. The reason for this proposition can be illustrated by the case of critiquing-based recommender systems: Critiquing-based recommender engines implement preference refinement as a process of further and further refining queries and narrowing down results. In order to do so, the values of an items attribute can be refined in each iteration of critiquing. This approach can be a strength and a weakness at the same time. It depends on the goal or context of use of the recommender system. If a user is searching for a specific item (e.g. something to purchase on an e-commerce platform) as in [2], the critiquing approach is promising. It helps the user in getting more and more specific results by narrowing down the result set w.r.t certain attributes. This can be accomplished, even if the user is not that familiar with the domain or start his search with a vague idea of how exactly the item will be. If, on the other hand, the user hopes for a variety of different, surprising items (e.g. in a music recommender), critiquing might be unsuitable to achieve this goal. Depending on the context of use, the concept of controlling-based recommenders can be more or less suited. In general, we would identify a user's goal in the recommendation process as a factor that could influence the appropriate type of control or recommendation concept to present to the user. Since this goal can change from session to session, it might be reasonable to allow user to change the default control method to a more convenient one during preference refinement.

Upon Result Display

During result display, one must keep in mind that any form of control that can be given to the user does not influence given recommendations as such. In the result display phase, the user gets to interact with the results of his query. We therefore divide the interaction into the following processes: (i.) *getting an overview* of the result set, (ii.) *inspecting the features* of a specific result, (iii.) *comparing items* (e.g. common/different attributes and values) (iv.) finding out *reasons for the recommendation* (v.) *choosing a result*. They have been identified based on processes in the related work:

(i.) *Overview:* The first one is obvious, since there must be a way to inspect all results produced during the process. It is further supported by Knijnenburg et al. [9] who accentuate the advantages of displaying the relationships between recommendation results and the users' configurations of the system [9] in an overview

(ii.) *Feature inspection*: The second is based on Pu and Chen [13]: They explain that supplementary information displayed with an item "may have significant impact on user satisfaction and confidence" [13]. This indicates that (ii.) is a recurring activity upon result display.

(iii.) *Result Comparison*: This phase must precede the choice of a result (v.)

(iv.) *Reasons for Recommendation*: Pu and Chen [13] also name *finding out the reasons for recommendation* (iv.) as a criterion for good interfaces. Also, the interfaces of SetFusion [12], SmallWorlds [4] or TasteWeights [1] explicitly put focus on visualizing how a recommendation was created, which also supports (iv.) as an important activity. Another indicator for the importance of reviewing this process upon result display.

(v.) *Choosing a result*: This directly follows from the recommender systems nature of being a "decision support tool" [13].

Any control method in the result display activity should be designed to support the user in at least one of the processes (i.) – (v.). We suppose that specifically for distrusting people, transparency of the recommendation process is an important issue. For users with a low level of trusting propensity it is more important to inspect how a result was created, than for other users. Consequently, more control w.r.t (iv.) is requested by distrusting users. One can also argue that users with high persistence would benefit from more control in finding out about the reasons for a recommendation. Generally, since maximizers are more likely to inspect and compare a great variety of results, it would make sense that they would request an increased level of control supporting all processes (i.) – (v.).

Design of control mechanisms: Regarding different levels of (domain) expertise, one should take into account the representation of controls as a factor which influences the amount of control requested by a user. We expect a carefully designed, explorable graphical representation of the recommendations to help the user in understanding how results were created. Furthermore, interacting with the results on a graphical interface could help users get better insight into the referred domain. This could reveal factors in the recommendation process which the user did not consider as relevant before. Such a learning process could result in an increase in domain expertise, and thus a further improvement of the configurations and of the recommendations. For novices, we specifically see the potential for such a process in a graphical user interface, which is why they could benefit from more control, under the precondition that the tools they use are easy to handle. The same holds for the graphical representation of the result set as such: Visual indicators such as links, colors and proximity for example are easier to read than text. They can convey meaning which the users may be unable to discover themselves without further support of the system. Especially novices in a domain might lack previous knowledge which is necessary to put given results into context. Verbert et al. [14] made an observation which supports this idea: TalkExplorer uses colors, connections and proximity in its result display, as explained above. Their results show that users gain more insight than when using a ranked list [14]. Parra and Brusilovski made another observation, which leads to a more general interpretation: they observed that native speakers tended to use less interactions than non-natives [12]. On a more general level, one could say that the graphical tools were easier to understand and use, than to accomplish the task on their own for people being at some kind of disadvantage. This disadvantage can be a language-barrier, lack of prerequisite knowledge, or other factors. Therefore, one could argue that users with disadvantages tend to rely more on controlling mechanisms in order to explore the domain and facilitate the task they are working on. Yet this is only the case if those control mechanisms are easy to understand and can be handled intuitively, thus reducing cognitive load. This idea is supported by the observations of Parra and Brusilovski [12]: they found that more items were selected from the results, when control methods (sliders and Venn diagram) were used on the result display. But they also found that the sliders were used more frequently than the diagram to narrow down the results. This supports the hypothesis that less ordinary mechanisms are used less often, at least upon the first encounter. We would attribute this to the fact that unknown graphical representations require cognitive effort, which is a sparse resource and needs to be spent on the resolution of the task itself. Experts, according to this reasoning, might not be so fond of many graphical tuning methods. As they understand the domain, they do not need as much support in interpreting the results they are presented with.

4 Conclusion

Throughout the previous sections we tried to explain, the claim that an increase in control is always appreciated by users of recommender systems cannot be accepted in its universality. Results from previous research showed that for specific user groups, more control does not necessarily result in higher user satisfaction. So far, recurring attributes that are used to differentiate between user groups are domain or system expertise, trusting propensity and persistence. By putting given results in the context of the three key activities of the recommendation process, preference elicitation, preference refinement and result display, it became clear that there are many cases in which the relationship between those user characteristics and the need for control remains unexamined. For those cases we offered suggestions and predictions based on more general research on control in recommender systems. Furthermore, for the three phases, additional factors could be identified that potentially distinguish further user groups: The nature of information required by the user and the context of use during preference elicitation, the interface in use and again the context or goal of use during preference refinement, and finally the design of control mechanisms upon result display. Those may also influence user satisfaction.

The guidelines formulated in this paper are merely of theoretical nature. However, they are based on an extensive review of existing work in the field. To empirically examine their validity could pose a starting point for new research to improve users' control in recommender systems.

5 Acknowledgements

The authors thank the German Research Council DFG for the friendly support of the research in the excellence cluster "Integrative Production Technology in High Wage Countries".

6 References

1. Bostandjiev, S., O'Donovan, J. and Höllerer, T. 2012. TasteWeights: A Visual Interactive Hybrid Recommender System. In *Proceedings of the sixth ACM conference on Recommender systems* (Dublin, Ireland, September 09 - 13, 2012). RecSys '12. ACM, New York, NY, 35-42. DOI=https://doi.org/10.1145/2365952.2365964
2. Chen, L. and Pu, P. 2006. Evaluating Critiquing-based Recommender Agents. In *Proceedings of the National Conference on Artificial Intelligence*. 21, 1, Menlo Park, CA, Cambridge, MA, London, AAAI Press, MIT Press, 1999, 2006, 157.
3. Ekstrand, M. D., Kluver, D., Harper, F. M. and Konstan, J. A. 2015. Letting Users Choose Recommender Algorithms: An Experimental Study. In *Proceedings of the 9th ACM Conference on Recommender Systems* (Vienna, Austria, September 16 - 20, 2015). RecSys '15. ACM New York, NY, 11-18. DOI= https://doi.org/10.1145/2792838.2800195
4. Gretarsson, B., O'Donovan, J., Bostandjiev, S., Hall, C. and Höllerer, T. 2010. Smallworlds: Visualizing Social Recommendations. In *Computer Graphics Forum*, 29, 3, Blackwell, 833-842. DOI=http://dx.doi.org/10.1111/j.1467-8659.2009.01679.x
5. Harper, F. M, Xu, F., Kaur, H., Condiff, K.; Chang, S. and Terveen, L.: Putting Users in Control of their Recommendations. In *Proceedings of the 9th ACM Conference on Recommender Systems (*Vienna, Austria, September 16 - 20, 2015). RecSys '15. ACM New York, NY, 3–10. DOI=https://doi.org/10.1145/2792838.2800179
6. Hijikata, Y., Kai, Y. and Nishida, S. 2012. The Relation Between User Intervention and User Satisfaction for Information Recommendation. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing* (Trento, Italy, March 26 - 30, 2012). SAC '12. ACM, New York, NY, 2002-2007. DOI= https://doi.org/10.1145/2245276.2232109
7. Jameson, A. and Schwarzkopf, E. 2002. Pros and Cons of Controllability: An Empirical Study. *In Proceedings of the Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems*. (May 29 – 31, 2002). AH '02.Springer-Verlag, London, UK, 193-202.
8. Knijnenburg, B. P., Reijmer, N. J. and Willemsen, M. C. 2011. Each to his Own: how Different Users Call for Different Interaction Methods in Recommender Systems. In *Proceedings of the fifth ACM conference on Recommender systems* (Chicago, Illinois, USA, October 23 – 27, 2011). RecSys '11. ACM, New York, NY, 141-148. DOI= https://doi.org/10.1145/2043932.2043960

9. Knijnenburg, B. P., Bostandjiev, S., O'Donovan, J. and Kobsa, A. 2012. Inspectability and Control in Social Recommenders. In *Proceedings of the sixth ACM conference on Recommender systems* (Dublin, Ireland, September 09 - 13, 2012). RecSys '12. ACM, New York, NY, 43-50. DOI= https://doi.org/10.1145/2365952.2365966
10. Loepp, B., Hussein, T. and Ziegler, J. 2014. Choice-based Preference Elicitation for Collaborative Filtering Recommender Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada, April 26 - May 01, 2014). CHI '14. ACM, New York, NY, 3085-3094. DOI= https://doi.org/10.1145/2556288.2557069
11. McNee, S. M., Lam, S. K., Konstan, J. A. and Riedl, J. 2003. Interfaces for Eliciting New User Preferences in Recommender Systems. In *Brusilovsky P., Corbett A. and de Rosis, F. (eds.): User Modeling 2003.* UM 2003. Lecture Notes in Computer Science, 2702, Springer, Berlin, Heidelberg, 178–187.
12. Parra, D. and Brusilovsky, P. 2015. User-controllable Personalization. A Case Study with SetFusion. In *International Journal of Human-Computer Studies,* 78, Elsevier, 43–67. DOI=10.1016/j.ijhcs.2015.01.007.
13. Pu, P., Chen, L. and Hu, R. 2012. Evaluating Recommender Systems from the User’s Perspective. Survey of the State of the Art. In *User Modeling and User-Adapted Interaction.*, 22, 4-5, Springer, 317–355. DOI=10.1007/s11257-011-9115-7
14. Verbert, K., Parra, D., Brusilovsky, P. and Duval, E. 2013. Visualizing Recommendations to Support Exploration, Transparency and Controllability. In *Proceedings of the 2013 international conference on Intelligent user interfaces* (Santa Monica, California, USA, March 19 - 22, 2013). IUI '13. ACM, New York, NY, 351-362. DOI= https://doi.org/10.1145/2449396.2449442
15. Sweller, J. 2010. Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. In *Educational Psychology Review,* 22, 2, Springer US, 123–138. DOI= 10.1007/s10648-010-9128-5
16. Valdez A.C, Ziefle, M., Verbert, K. 2016. HCI for Recommender Systems: the Past, the Present and the Future. In *Proceedings of the 10th ACM Conference on Recommender Systems* (Boston, Massachusetts, USA, September 15 - 19, 2016). RecSys '16. ACM, New York, NY, 123-126. DOI=https://doi.org/10.1145/2959100.2959158