# Each to His Own: How Different Users Call for Different Interaction Methods in Recommender Systems

Bart P. Knijnenburg
Department of Informatics
University of California, Irvine
Bart.K@uci.edu

Niels J.M. Reijmer
Human-Technology Interaction group
Eindhoven University of Technology
N.J.M.Reijmer@student.tue.nl

Martijn C. Willemsen
Human-Technology Interaction group
Eindhoven University of Technology
M.C.Willemsen@tue.nl

## ABSTRACT

This paper compares five different ways of interacting with an attribute-based recommender system and shows that different types of users prefer different interaction methods. In an online experiment with an energy-saving recommender system the interaction methods are compared in terms of perceived control, understandability, trust in the system, user interface satisfaction, system effectiveness and choice satisfaction. The comparison takes into account several user characteristics, namely domain knowledge, trusting propensity and persistence. The results show that most users (and particularly domain experts) are most satisfied with a hybrid recommender that combines implicit and explicit preference elicitation, but that novices and maximizers seem to benefit more from a non-personalized recommender that just displays the most popular items.

## Categories and Subject Descriptors

H.1.2. **[Models and principles]**: User/Machine Systems–*software psychology*; H.4.2. **[Information Systems Applications]**: Types of Systems–*decision support*; H.5.2 **[Information Interfaces and Presentation]**: User Interfaces–*evaluation/methodology, interaction styles, user centered design*

## General Terms

Measurement, Design, Experimentation, Human Factors, Theory.

## Keywords

Recommender systems, human-computer interaction, usability, user experience, user interfaces, preference elicitation methods

## 1. INTRODUCTION

This paper investigates how different types of users react to recommender systems that employ a wide range of interaction methods. The way users interact with a recommender system seems to have an impact on their satisfaction with the system [20,28]. But whereas recommender systems typically employ the same interaction method for all users, the appreciation for that method may not be universal. In fact, researchers have claimed that recommender systems should move beyond their "one-size-fits-all" approach and adapt their interaction method to the user [19].


*RecSys'11*, October 23–27, 2011, Chicago, Illinois, USA.


Our earlier work showed that domain experts and novices differed in their preferred interaction method [12,13]. That work compared two types of explicit preference elicitation for an attribute-based recommender system. In the current paper we consider explicit, implicit and hybrid preference elicitation methods, but we also include a fixed TopN and a sortable table in our comparison; these interaction methods are in fact much simpler and non-personalized. Moreover, aside from domain knowledge, we consider how the best interaction method may differ for users with different levels of persistence and trusting propensity.

## 2. THEORY

The current paper considers attribute-based recommender systems: systems that base their recommendations on the values of the attributes on which each item in the system is defined.

### 2.1 Decision Strategies

A good recommender system suggests items that match user preferences. Research suggests that preferences are not static, predefined dispositions, but are in fact constructed whenever a decision needs to be made [6]. The construction of preferences is not a straightforward endeavor: there are several distinct decision strategies that can be employed to decide what to choose [4]. Specifically, Bettman, Luce and Payne discuss the following strategies[1]:

- Weighted adding: Like in multi-attribute utility theory, the consumer assigns a weight Wj to each attribute j, and sums the product of the weight times the value of each attribute $\Sigma(Vij * Wj)$ for each item i. She then chooses the item with the highest outcome. The weighted adding strategy is compensatory: if an item has a low value on one attribute, this can be compensated by a high value on another attribute.
- Satisficing: The consumer evaluates items one by one. If an item meets a certain threshold for each of its attributes, the consumer chooses that item. Satisficing is a non-compensatory strategy: if an attribute value does not meet the threshold, the item will not be considered by the user, regardless of the values of the other attributes.
- Lexicographic: The consumer chooses the item with the highest value on the most important attribute. If there are several items with the highest value, the second-most important attribute is considered. The lexicographic strategy is non-compensatory in the other direction: the item with the highest value on the most important attribute is selected, regardless of the other attribute values.

[1] Bettman, Luce and Payne propose a total of 8 strategies; we discuss only those that pertain to the rest of the discussion.

Traditionally, websites have either filtered the list of items based on minimum attribute values provided by the user (implementing the satisficing strategy), or allowed the user to sort items by attribute values (implementing the lexicographic strategy). Most attribute-based recommender systems, on the other hand, implement the weighted adding strategy. Because this strategy requires extensive calculation, especially when the number of items and/or attributes is large, automating it can potentially reduce a large cognitive burden from the user.

Bettman, Luce and Payne argue that the decision strategy selected by the consumer critically depends on her personal characteristics [4], and researchers have shown that users prefer a recommender system that uses a decision strategy similar to the one they would use themselves [1,2]. In effect, users with different characteristics (who use different strategies) may prefer different interaction methods. On the other hand, one may argue that regardless of the users' preferred strategy, a recommender system that implements a compensatory strategy leads to normatively superior decisions [25]. This leads to an interesting divergence: the *process* (i.e. using the system, [18]) may be most satisfying when the interaction method matches the user's decision strategy, while the *outcome* (i.e. chosen items, [17]) may be most satisfying when the method is compensatory.

Below we describe five interaction methods, each supporting a different decision strategy. Consequently, we describe three user characteristics, and reason which interaction method(s) would be preferred by users with different levels of these characteristics, in terms of both process and outcome.

## 2.2 Five Interaction Methods

We constructed five interaction methods that have previously been employed in both commercial and research recommender systems, and cover a wide array of supported decision strategies.

### *2.2.1 TopN*

Our TopN method provides recommendations that are the same for each user (non-personalized) and ranked by overall popularity of the items (based on decision logs from earlier experiments). In TopN, users do not have to decide on attribute weights. TopN is our baseline comparison interaction method: there is virtually no interaction, as the user cannot change the order of the recommendations. Personalized recommendations are typically more accurate than non-personalized recommendations, and researchers in the field of recommender systems typically assume that they are therefore also more satisfying.

### *2.2.2 Sort*

In the Sort method, users may sort recommendations on any attribute (sorting is done in descending order of attractiveness). The initial ordering is on popularity, but once another attribute has been chosen, the user cannot return to this ordering[2]. Sort implements a lexicographic strategy, which requires that users know their most important attribute. Aside from this user-selected sort-attribute, this system is non-personalized. Users may change the sort-attribute as many times as they want during the interaction.

[2] The option to sort on popularity is not available because it is not an attribute in the dataset. Initially sorting on any other attribute (or alphabetically) would however give this method a clear disadvantage.

### *2.2.3 Explicit*

The Explicit method implements the weighted adding strategy. Users can directly set their preferences by indicating the weight they assign to each of the attributes. For each alternative, these weights are then multiplied by the (normalized) attribute values, and summed to get a utility score. Alternatives are then ranked by utility score, and the best ones are recommended. Explicit requires that users make tradeoffs between the attribute weights.

### *2.2.4 Implicit*

The Implicit method also employs a multi-attribute utility calculation, but it automatically determines the attribute weights based on the user's browsing behavior. Unlike the Explicit and Hybrid version, the weights are not displayed in this version. Whenever users preview or select a recommended item, a set of rules analyzes this behavior and updates attribute weights accordingly. The rules are based on the analysis of browsing and preference logs from previous versions of the system. The implicit system does not require users to make tradeoffs between attribute weights.

### *2.2.5 Hybrid*

The Hybrid method combines the Explicit and Implicit approaches by automatically updating the attribute weights while at the same time offering users the option to change the weights themselves. The simultaneous influence of manual and automatic updates on the weights adds some complexity to this method. However, it combines the convenience of automatic preference elicitation with the option to monitor and control the weights explicitly. The Implicit and Hybrid methods are further explained in [22].

## 2.3 User Characteristics

We do not argue that any of the interaction methods described in the previous section is generally better than the others. Instead, as decision strategy selection depends on specific user characteristics, we hypothesize that the preferred interaction method also depends on these characteristics. Literature suggests three user characteristics that may have an impact on strategy selection: domain knowledge, trusting propensity and persistence. However, this literature does not cover the link from strategy selection to interaction methods: this part of our work is exploratory in nature.

### *2.3.1 Domain Knowledge*

Bettman, Luce and Payne argue that the preferred decision strategy may depend on users' expertise in the choice domain [4]. Specifically, methods like weighted adding allow experts to make use of their intimate knowledge of the product attributes. In principle, experts' knowledge of attributes makes them better equipped to use a personalized attribute-based recommender system, leading to better outcomes. However, Kamis and Davern [11] show that experts perceive personalized recommenders as less useful than novices. They argue that such systems may take away their control over their decisions (see also [7,25]). Not much is known about the perception of control in recommender systems. Implicit recommender systems may be most susceptible to this problem [10], and may therefore result in a less satisfying process. Explicit and Hybrid provide control over the attribute weights, and may thus be the optimal for expert users in terms of both process and outcome [12,13].

Novices typically lack attribute knowledge [3,5], which may prohibit them to effectively use a personalized attribute-based recommender system that leverages such knowledge (i.e. Explicit and Hybrid) [15,19,21]. They may on the other hand be more satisfied using a strategy that does not require this knowledge, such as a lexicographic strategy (i.e. the Sort method), or any method that does not require knowledge of attributes (i.e. Implicit and TopN) [6,21]. The Implicit method has the added benefit that it uses a personalized compensatory strategy, which should result in better outcomes [25]. However, Kramer shows that domain novices prefer simple, transparent interaction methods [16]. They may thus be more satisfied when using TopN and Sort (as compared to Explicit, Implicit and Hybrid), also because these methods are more common on current websites.

### *2.3.2 Trusting Propensity*

A recommender system is in essence a persuasive system [8]; it tries to persuade its users to follow its recommendations. Komiak and Benbasat show that users' trust in the system influences how well this persuasion works: only users who trust the system will continue to use it [15]. Trust in the system can be developed over an extended period of interaction with a recommender system. However, in the initial interaction period, trust in the system is mainly influenced by users' general trusting propensity [27].

Unless they understand how the system works [9], distrusting users usually want to take more control over the system [26]. If this is not possible, it will lower their satisfaction or even cause reactance (actively countering the system's advice) [7]. Distrusting users may therefore like systems that are easy to understand (i.e. TopN and Sort) and in which they can control the way it represents their preferences (i.e. Explicit and Hybrid). The Implicit method may thus be the least appealing (in terms of process, but not necessarily in terms of outcome) for distrusting users.

### *2.3.3 Persistence*

Schwartz [23] describes a prominent distinction in decision-makers between satisficers and maximizers. Satisficers and maximizers differ in their level of persistence when making a decision: Satisficers will stop the decision-making process when they encounter an item that meets their minimal criteria, while maximizers aspire the best possible option [24]. For satisficers any interaction method may suffice, but Implicit may result in the best outcomes, because the system updates the recommendations to provide similar items as soon as the user selects the first item.

Maximizers usually engage in extensive product comparison and more counterfactual thinking than satisficers [23]. Because maximizers always consider the possibility that there could be a better option than the one they chose, they will anticipate more postdecision regret, and their choice satisfaction will therefore be lower than for satisficers. The effects of counterfactual thinking are aggravated when the used decision strategy is compensatory, because in those cases making the decision involves a (reversible) tradeoff, leading to even more anticipated postdecision regret [23]. Explicit and Hybrid implement a compensatory strategy, and tradeoffs in these systems are reversible as users have explicit control over their preference weights. Moreover, despite the fact that Sort is non-compensatory, tradeoffs may be most palpable in this system, because the user has to make an "all or nothing" tradeoff where every other attribute is sacrificed to get a better value of the sorted attribute. Explicit, Sort and Hybrid may thus lead to lower choice satisfaction for maximizers.

## 3. ONLINE STUDY

To test our main hypothesis that the preferred interaction method depends on user characteristics, we performed an online user experiment with a recommender system for energy-saving measures. The system, an updated version of the one used in [12] and [13], selects recommendations from 80 energy-saving measures, ranging from behavioral solutions ("turning off the lights when you leave the room") to product- or service-based solutions ("installing CFLs"). Each measure is defined on 8 attributes: initial effort, continuous effort, initial costs, savings in Euro/year, savings in kWh/year, return on investment, overall environmental effects, and comfort. The domain of energy saving was used because of its relevance in today's society. Also, in this domain it is natural to select multiple measures, and the choice-situation is more realistic than fake online shopping.

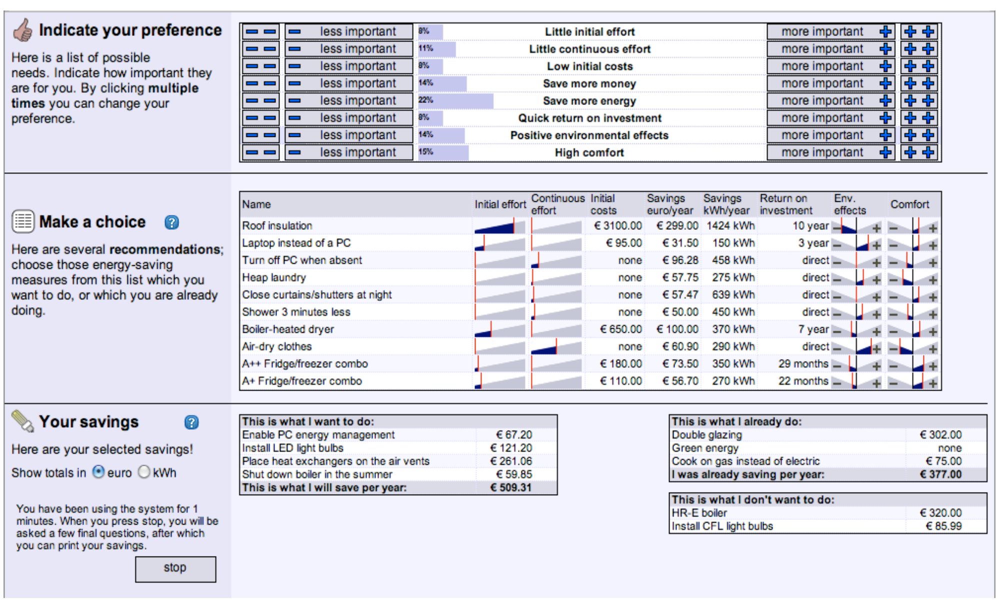


**Figure 1. The online recommender system for energy-saving measures. The interface is translated from Dutch. The Explicit interaction method is shown (which looks identical to the Hybrid method); the TopN, Sort and Implicit method are similar except that these do not display any attribute weights in the top part of the interface.**

The system is displayed in Figure 1. The middle part shows the first 10 recommendations. In the TopN system these have a static order. In the Sort system, users can change the recommendations by sorting on an attribute (by clicking on its name in the top row). In the Explicit and Hybrid systems, users can inspect and change the attribute weights in the top part of the system. Changing the weights immediately updates the recommendations. In the Hybrid and Implicit system, the system itself can change the weights (and therefore the recommendations) automatically.

When the user clicks on a recommendation, the system shows some information about the energy saving measure, and allows the user to choose what to do with it: “I don’t know yet” (default), “I want to do this”, “I’m already doing this”, and “I don’t want to do this” (Figure 2). For the latter three options, three separate lists are shown at the bottom part of the system. When the user puts a recommendation in one of these lists, it is removed from the list of recommendations, and a new recommendation is added to the bottom of the list. This way users can potentially browse through all recommendations, even when they do not have a mechanism to change the recommendations (as is the case for TopN).

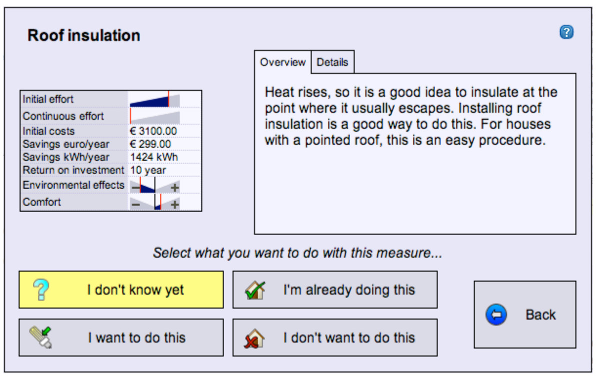


**Figure 2. Screen shown to users when they click on an item**

158 participants were recruited by an external company to participate in the experiment. 11 participants were removed as they had spend too little time with the system (< 2.5 minutes). Of the remaining 147 participants 79 (54%) were male, their age ranged from 13-77 with an average of 40.0 and a standard deviation of 15.9. There were 29 students, 93 working and 25 retired participants. 23 participants finished high school, 24 had an intermediate degree, 53 a bachelor’s degree, and 47 a master’s degree. 27 participants received the Top-N, 30 the Sort, 29 the Explicit, 28 the Implicit, and 33 the Hybrid system.

Participants were first asked 7 demographics questions and 18 questions (statements to which the user could agree or disagree on a 5-point scale) to determine their level of domain knowledge, trusting propensity and persistence. Consequently they were randomly assigned to one of the 5 interaction methods, and instructed on how to use the system. They used the system for as long as they liked. Afterwards, participants were asked 44 questions (mostly statements to which the user could agree or disagree on a 5-point scale) about their experience of using the system. The full set of questionnaires can be found in [22].

Finally, participants were given the option to enter in a raffle for either a 20 Euro gift certificate or an energy-saving light bulb worth 25 Euro (one prize was awarded to every 10th participant). Upon leaving the experiment, users were offered the opportunity to print the measures they selected using the system.

## 3.1 Questionnaires

The pre-experimental questionnaire was submitted to an exploratory factor analysis. After removing one item with low communality, the expected three-factor solution provided an adequate fit, measuring domain knowledge (7 items), trust inclination (6 items) and persistence (4 items). Because the analysis of our results considers an interaction between user characteristics and interaction methods, the factor scores were saved as standardized linear scale variables.

The post-experimental questionnaire covered several relevant aspects of the *process* of using the system (control, understandability, trust in the system, and five general items from the Questionnaire of User Interface Satisfaction, or QUIS), as well as the *outcome* of using the system (choice satisfaction). The QUIS scores (9-point scales) were simply summed, while the other items were submitted to an exploratory factor analysis. After removing 1 item with low communality and 10 items with high cross-loadings, the expected five-factor solution provided an adequate fit, measuring control (7 items), understandability (8 items), trust in the system (4 items), perceived system usefulness (5 items) and choice satisfaction (4 items). These standardized factors are used as outcomes in the analysis of our results.

## 3.2 Results

### *3.2.1 Main effects of interaction methods*

Although the five interaction methods differ substantially, none of them turns out to be the overall winner. As expected, there are, on average, no significant differences between interaction methods on the outcome measures: perceived control [$p = .874$], understandability [$p = .329$], trust in the system [$p = .840$], interface satisfaction (QUIS) [$p = .140$], perceived effectiveness [$p = .142$], or satisfaction with the chosen measures [$p = .525$][3]. Instead, we find that the best interaction method depends on the characteristics of the user. Below, we describe the significant interactions between interaction methods and user characteristics.

### *3.2.2 Domain knowledge*

We reasoned that the main difference between experts and novices would be their level of attribute knowledge, and that this would influence their preferred interaction method. Accordingly, we hypothesized that novices prefer methods that do not require intimate attribute knowledge (TopN, Sort and Implicit, although the latter may not be transparent enough), while experts prefer methods that give them the control needed to leverage such knowledge (Explicit and Hybrid).

The effect of domain knowledge on **perceived control** differs marginally significantly between interaction methods [overall $p < .1$][4]. As hypothesized, there is a significant negative relation between domain knowledge and perceived control in the TopN

[3] All main effects are tested by regressing each outcome variable on the user characteristics and the interaction methods. Our results are robust against progressively deleting outliers (up to 10% of the data).

[4] Interaction effects are tested by regressing each outcome variable on the interaction method and a separate effect of each characteristic for each interaction method. We first test whether the effect of the characteristic is the same for each interaction method (overall test), and then test the effect per interaction method separately.

system [p < .05]. Novices thus perceive most control in this system, while experts perceive least control in this system. Contrary to Kamis and Davern's [11] results, control is not lower for experts in other interaction methods, not even in the Implicit method. We ourselves hypothesized that for experts, the perceived control would be highest in the Explicit and Hybrid systems (see Figure 3), but these effects are not significant.

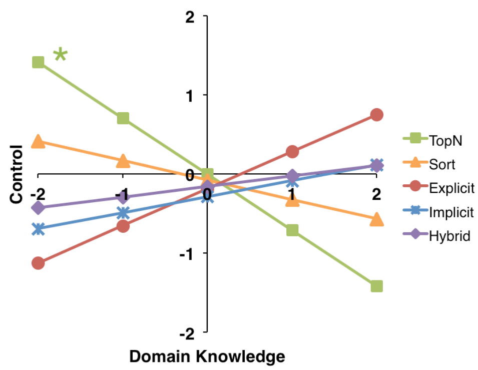


**Figure 3. Effect of domain knowledge on perceived control for different interaction methods. Marks indicate significant slopes: [1] p < .1, * p < .05, ** p < .01, *** p < .005, **** p < .001. The scales on the axes in sample standard deviations.**

The **understandability** of the system is higher for domain experts regardless of the interaction method [p < .05]. In general, this effect does not differ significantly between systems [overall p = .563], but Figure 4 and the analysis of the interaction effects show that the effect of domain knowledge is only significant in the Hybrid system [p < .05]. Experts thus seem to understand the Hybrid system the best, while novices understand this system the least.

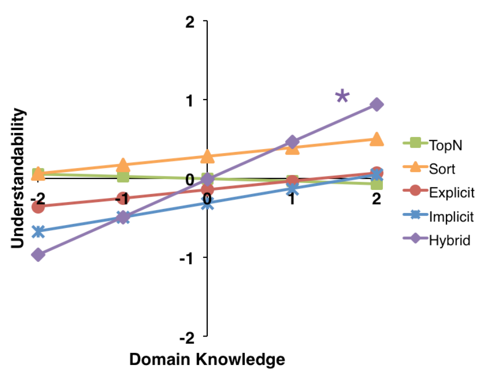


**Figure 4. Effect of domain knowledge on the understandability of different interaction methods**

The effect of domain knowledge on **user interface satisfaction** (QUIS) differs significantly between interaction methods [overall p < .005]. As hypothesized, there is a significant positive relation in the Hybrid system [p < .005], indicating that novices are least satisfied with the Hybrid user interface, while experts are most satisfied with this interface (Figure 5). Interestingly, this effect cannot be explained by a higher level of control for experts in this system, but it may be due to a higher understandability. There is also a significant negative relation between domain knowledge and user interface satisfaction in the Explicit system [p < .01], indicating (contrary to our expectations) that experts are in fact least satisfied with the user interface of this interaction method. The other interaction effects are not significant: specifically, the TopN, Implicit and Sort interfaces are no more satisfying for novices than for experts.

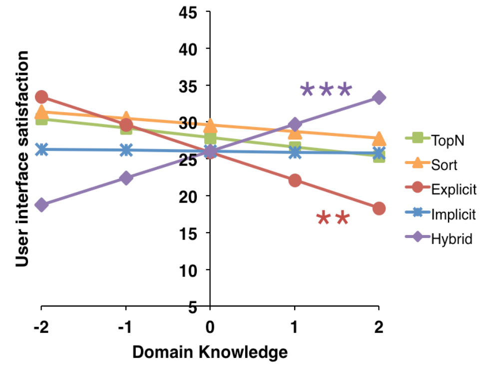


**Figure 5. Effect of domain knowledge on user interface satisfaction for different interaction methods**

The effect of domain knowledge on **perceived system effectiveness** also differs significantly between interaction methods [overall p < .001]. As hypothesized, there is a significant negative relation between domain knowledge and perceived system effectiveness in the TopN system [p < .005] and a marginally significant negative relation in the Sort system [p < .1]. Also as hypothesized, there is a significant positive relation in the Hybrid system [p < .01]. Accordingly, Figure 6 shows that experts find the Hybrid system most effective, while novices find the TopN Sort systems most effective. The effects of domain knowledge in the Explicit and Implicit conditions are not significant.

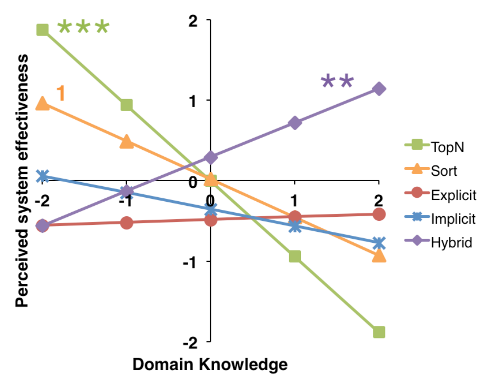


**Figure 6. Effect of domain knowledge on perceived system effectiveness for different interaction methods**

Finally, **choice satisfaction** significantly increases with domain knowledge [p < .005]. In general, this effect does not differ per interaction method [overall p = .175], but Figure 7 and the analysis of the interaction effects show that, in line with our expectations, the effect of domain knowledge is only significant in the Explicit [p < .05] and Hybrid [p < .01] conditions. Experts are thus more satisfied with their choices, but only in these conditions. The Implicit condition, which we expected to lead to good outcomes for experts and novices alike, is marginally better for experts [p < .1]. For TopN, the effect seems to be decreasing as hypothesized, but this effect is not significant [p = .342]. The effect for Sort is also not significant.

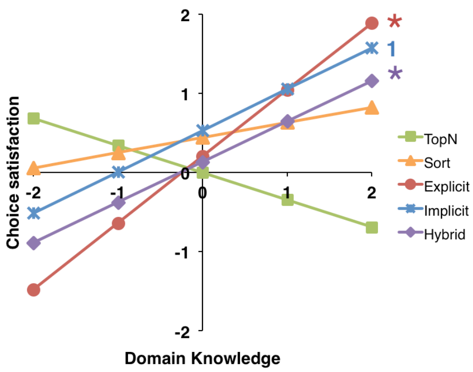

**Figure 7. Effect of domain knowledge on choice satisfaction for different interaction methods**

### 3.2.3 *Trusting propensity*

We reasoned that distrusting users would only use systems that do not require a high level of trust, and would therefore prefer interaction methods that are easy to understand (TopN and Sort) or easy to control (Explicit and Hybrid). However, we found no main differences between the interaction in terms of control and understandability (see section 3.2.1), and also no interaction effects between interaction methods and trusting propensity [overall p = .434 and .760 respectively]. This means that whatever effects of trusting propensity we find, these cannot be explained by differences in perceived control or understandability.

As expected, **Trust in the system** increases with trusting propensity [p < .05]. In general, this effect does not differ significantly between systems [p = .105], but Figure 8 and the analysis of the interaction effects show that this effect is mainly driven by the users in the Sort condition [p < .01]. Contrary to our expectations, the Sort system inspires the most trust in people with a high trusting propensity, but it inspires the least trust in people with a low trusting propensity. The other specific effects are not significant.

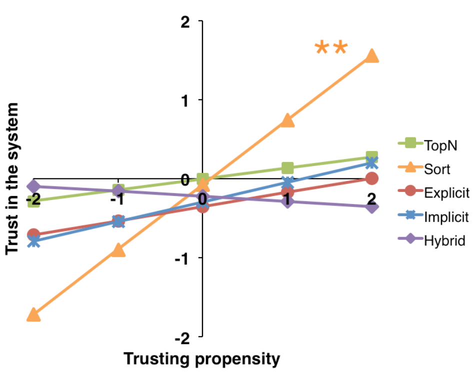

**Figure 8. Effect of trusting propensity on trust in the system for different interaction methods**

**User interface satisfaction** (QUIS) marginally significantly increases with trusting propensity [p < .1], and this effect differs significantly between interaction methods [overall p < .05]. Specifically, there is a significant positive relation between trusting propensity and user interface satisfaction in the Explicit system [p < .001] and the Implicit system [p < .05], and a somewhat significant positive relation in the TopN system [p = .10]. Figure 9 shows that users with a lower trusting propensity have a lower interface satisfaction in the Explicit, Implicit and TopN conditions; the effect for Implicit is in line with our expectations, but the effects for Explicit and TopN are not. The effects of trusting propensity in the Sort and Hybrid conditions are not significant.

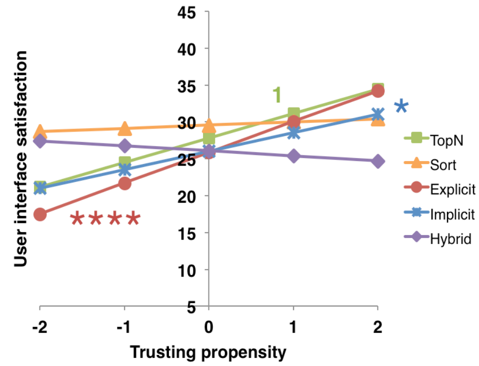

**Figure 9. Effect of trusting propensity on user interface satisfaction for different interaction methods**

The effects of trusting propensity on **perceived system effectiveness** are similar to those of trusting propensity on user interface satisfaction: there is a significant difference between interaction methods [overall p < .05], an increase in the Explicit [p < .05] and Implicit [p < .05] conditions, and a marginal increase in the TopN condition [p < .1] (see Figure 10). Again, these effects are in contrast with our expectations for Explicit and TopN, but in line with our expectations for Implicit.

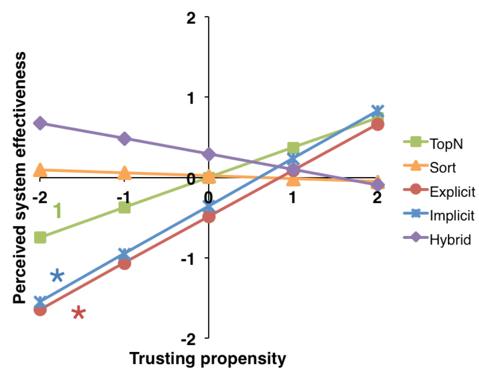

**Figure 10. Effect of trusting propensity on perceived system effectiveness for different interaction methods**

In line with our expectation that trusting propensity affects process but not outcome, **choice satisfaction** is not related to trusting propensity [p = .701], and the effect does not differ per interaction method [p = .791].

### 3.2.4 *Persistence*

There are no differences between satisficers and maximizers in terms of control [p = .524], understandability [p = .424], interface satisfaction [p = .755], or system effectiveness [p = .859], which confirms our expectation that persistence affect outcome but not process. We reasoned that in general maximizers would experience an increased level of anticipated postdecision regret. Interestingly, **choice satisfaction** significantly increases with persistence [p < .005]. There is no significant difference in the effect of persistence on choice satisfaction between the different interaction methods [p = .278], but Figure 11 and the analysis of the interaction effects show that the effect of persistence is only significant in the TopN condition [p < .01]. Arguably, this may be due to the lack of tradeoffs required to use the TopN system.

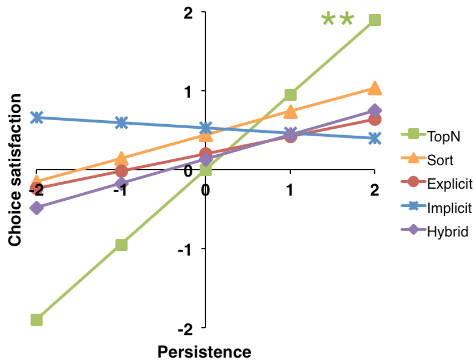


**Figure 11. Effect of persistence on choice satisfaction for different interaction methods**

## 3.3 Discussion

Despite the differences between the interaction methods, none of them turns out to be an overall winner. Instead, we show that the "best" interaction method depends on several user characteristics.

### *3.3.1 Domain knowledge*

As hypothesized, novices seem to prefer interaction methods that do not require intimate knowledge of attributes. Especially when using the TopN system, they feel more in control (see Figure 3), and they perceive an increased effectiveness (see Figure 6) and choice satisfaction (see Figure 7). The Sort interface comes in second place on most measures. The Implicit interaction method is arguably not transparent enough to be satisfying for novices. Notably, novices rate the understandability low for all systems.

We predicted that experts would prefer the Explicit and Hybrid systems, which allow them to control the system and leverage their attribute knowledge. These methods indeed lead to a higher choice satisfaction for experts. Experts seem to perceive most control in the Explicit system (although this effect is not significant; see Figure 3), but because they are not satisfied with the Explicit user interface (see Figure 5), they only perceive the Hybrid system as significantly more effective (see Figure 6). The combination of the superior control of the Explicit system and the convenience of the Implicit system seems to be best for experts.

### *3.3.2 Trusting propensity*

We expected that distrusting users would prefer to use systems that are easy to understand (TopN and Sort) or offer more perceived control (Explicit and Hybrid); hence Implicit would be the worst system for distrusting users. Indeed, distrusting users are not satisfied with this system or its user interface (see Figures 9 and 10). But contrary to our hypothesis, distrusting users are equally unsatisfied with the TopN and Explicit systems, and moreover seem to particularly mistrust the Sort system (see Figure 8). However, as the hypotheses about perceived system effectiveness for distrusting individuals were based on hypothesized differences in control and understandability, the apparent lack of such differences make our initial hypotheses unfounded.

### *3.3.3 Persistence*

Contrary to Schwartz's theory, in our experiment maximizers are more satisfied with their choices than satisficers. It may be that any system that collects detailed information about 80 energy-saving measures increases maximizers' potential to select better measures. That said, the higher satisfaction for maximizers is only significant for the TopN interaction method (see Figure 11). A possible explanation for this could be that tradeoffs are avoided in the TopN system, thereby avoiding counterfactual thinking and thus anticipated post-decision regret.

# 4. CONCLUSION AND FUTURE WORK

In our experiment we have shown that the best interaction method for an attribute-based recommender system depends on the characteristics of the user. This has several implications for the design of such systems.

In general, it seems to be a good idea to combine explicit and implicit preference elicitation in a hybrid recommender system. In our experiment, the Hybrid system dominated the Explicit and Implicit systems in nearly every aspect. Specifically, the Hybrid system works best for satisficers who quickly want to attain satisfactory results, as well as for experts who can tweak their preferences to get superior recommendations. The Hybrid system also doesn't seem to endure a negative reaction from distrusting users. However, a simple TopN system may be preferred in some cases: a hybrid recommender may be too complex for novices, and a non-personalized approach may also be more suitable for maximizers who experience more regret when making tradeoffs.

Designers seem to have to find a way to combine the simplest method (TopN) and the most complex method (Hybrid), while avoiding their respective downsides. Combining these methods in a single recommender system is a difficult design challenge. One option is to spatially separate the two interaction methods in different sections of the system. Another option is to temporally separate them: start with the TopN, carefully introduce implicit recommendations, and then introduce explicit controls as well. A final option is assign the correct method to each user: try to discover before (or during) the interaction what the user's characteristics are, and then tailor the interface to her specific needs.

More generally, we show that designers and researchers alike should investigate the impact of their system on the user's satisfaction in terms of both *process* and *outcome*. Being satisfied with the system itself and the outcomes of using it are two separate concerns, that may at times even be in conflict with one another.

Our study was conducted with a relatively small sample of users (for an experiment with 5 conditions) from a somewhat heterogeneous population. Specifically, we found a strong positive correlation in our sample between expertise and persistence, and this may have reduced the power of our analyses. Moreover, when estimating the effects of the user characteristics for each of the interaction methods separately, each of these tests is based on a mere 28-33 participants per condition. This reduces the power of our analyses to such an extent that it becomes hard to test some of the (more subtle) effects that seem to exist in the different graphs. In other words, a non-significant effect does not mean that it is unquestionably absent, but rather that it may be too small to be reliably tested in our limited sample. The domain, which encourages multiple decisions, may also have dampened the effects. Moreover, our results pertain to attribute-based recommender systems, and we are hesitant to extend them beyond this scope. Specifically, we cannot draw any conclusions about the effect of user characteristics on collaborative filtering recommender systems. We provide an initial comparison of explicit and implicit preference elicitation methods for collaborative filtering systems in [14], but future work should investigate whether here too the preferred interaction method is significantly different for different user characteristics.

Despite these drawbacks, we feel confident to argue that the user experience of a recommender system critically depends on its interaction method, and that a careful consideration of user characteristics in the design and evaluation of these methods can lead to significantly better systems.

## 5. ACKNOWLEDGEMENTS

We thank Alfred Kobsa and Robin Keller for their extensive feedback on several early drafts of this paper.

## 6. REFERENCES

[1] Aksoy, L. 2006. Should Recommendation Agents Think Like People? *Journal of Service Research*. 8, 4, 297-315. DOI= http://dx.doi.org/10.1177/1094670506286326

[2] Al-Natour, S. et al. 2006. The role of design characteristics in shaping perceptions of similarity: The case of online shopping assistants. *J. Assoc. for Inform. Syst*. 7, 12, 821-861.

[3] Alba, J.W. and Hutchinson, J.W. 1987. Dimensions of Consumer Expertise. *Journal of Consumer Research*. 13, 4, 411-422. DOI= http://dx.doi.org/10.1086/209080

[4] Bettman, J.R. et al. 1998. Constructive consumer choice processes. *Journal of consumer research*. 25, 3, 187-217. DOI= http://dx.doi.org/10.1086/209535

[5] Chernev, A. 2003. When More Is Less and Less Is More: The Role of Ideal Point Availability and Assortment in Consumer Choice. *Journal of Consumer Research*. 30, 2, 170-183. DOI= http://dx.doi.org/10.1086/376808

[6] Coupey, E. et al. 1998. Product Category Familiarity and Preference Construction. *Journal of Consumer Research*. 24, 4, 459-468. DOI= http://dx.doi.org/10.1086/209521

[7] Fitzsimons, G.J. and Lehmann, D.R. 2004. Reactance to Recommendations: When Unsolicited Advice Yields Contrary Responses. *Marketing Science*. 23, 1, 82-94. DOI= http://dx.doi.org/10.1287/mksc.1030.0033

[8] Fogg, B. 2003. *Persuasive technology : using computers to change what we think and do*. Morgan Kaufmann Publishers.

[9] Jameson, A. 2003. Adaptive interfaces and agents. *Human-computer interaction handbook*. J. A. Jacko and A. Sears, eds. Laurence Earlbaum Ass., New York, NY. 305-330.

[10] Jones, N. et al. 2009. How Users Perceive and Appraise Personalized Recommendations. *User Modeling, Adaptation, and Personalization*. G.-J. Houben et al., eds. Springer, Heidelberg, Germany. 461-466. DOI= http://dx.doi.org/10.1007/978-3-642-02247-0_53

[11] Kamis, A. and Davern, M.J. 2004. Personalizing to product category knowledge: exploring the mediating effect of shopping tools on decision confidence. *Proc. of the 37th Annual Hawaii International Conference on System Sciences.* 10 pp. DOI= http://dx.doi.org/10.1109/HICSS.2004.1265476

[12] Knijnenburg, B.P. and Willemsen, M.C. 2010. The effect of preference elicitation methods on the user experience of a recommender system. *Extended abstracts on Human factors in computing systems*. ACM, New York, NY, 3457-3462. DOI= http://doi.acm.org/10.1145/1753846.1754001

[13] Knijnenburg, B.P. and Willemsen, M.C. 2009. Understanding the effect of adaptive preference elicitation methods on user satisfaction of a recommender system. *Proc. of the third ACM conf. on Recommender systems*. ACM, New York, NY, 381-384. DOI= http://doi.acm.org/10.1145/1639714.1639793

[14] Knijnenburg, B.P. et al. Explaining the user experience of recommender systems. Accepted to *User Modeling and User-Adapted Interaction*. http://t.co/cC5qPr9

[15] Komiak, S.Y.X. and Benbasat, I. 2006. The effects of personalization and familiarity on trust and adoption of recommendation agents. *Mis Quarterly*. 30, 4, 941–960.

[16] Kramer, T. 2007. The Effect of Measurement Task Transparency on Preference Construction and Evaluations of Personalized Recommendations. *J. of Marketing Research*. 44, 2, 224-233. DOI= http://dx.doi.org/10.1509/jmkr.44.2.224

[17] McNee, S.M. et al. 2002. On the recommending of citations for research papers. *Proc. of the 2002 ACM conf. on Computer Supported Cooperative Work*. ACM, New York, NY, 116-125. DOI= http://doi.acm.org/10.1145/587078.587096

[18] McNee, S.M. et al. 2003. Interfaces for Eliciting New User Preferences in Recommender Systems. *User Modeling 2003*. P. Brusilovsky et al., eds. Springer, Heidelberg, Germany. 178-187. DOI=http://dx.doi.org/ 10.1007/3-540-44963-9_24

[19] Pereira, R.E. 2000. Optimizing human-computer interaction for the electronic commerce environment. *Journal of Electronic Commerce Research*. 1, 1, 23-44.

[20] Pu, P. et al. 2008. Evaluating product search and recommender systems for E-commerce environments. *Electronic Commerce Research*. 8, 1-2, 1-27. DOI= http://dx.doi.org/10.1007/s10660-008-9015-z

[21] Randall, T. et al. 2007. User Design of Customized Products. *Marketing Science*. 26, 2, 268-280. DOI= http://dx.doi.org/10.1287/mksc.1050.0116

[22] Reijmer, N.J.M. Willemsen, M.C. and Knijnenburg, B.P. The effect of user characteristics on the satisfaction of different interaction methods for recommender systems. Technical report. http://bit.ly/ofPenD

[23] Schwartz, B. 2004. *The paradox of choice why more is less*. HarperCollins, New York, NY.

[24] Schwartz, B. et al. 2002. Maximizing versus satisficing: Happiness is a matter of choice. *Journal of Personality and Social Psychology*. 83, 5, 1178-1197. DOI= http://dx.doi.org/10.1037//0022-3514.83.5.1178

[25] Spiekermann, S. and Paraschiv, C. 2002. Motivating human–agent interaction: Transferring insights from behavioral marketing to interface design. *Electronic Commerce Res*. 2, 3, 255-285. DOI= http://dx.doi.org/10.1023/A:1016062632182

[26] Vries, P. 2004. *Trust in systems: effects of direct and indirect information*. PhD thesis. Technische Universiteit Eindhoven.

[27] Wang, W. and Benbasat, I. 2007. Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs. *J. of Manage. Inform. Syst*. 23, 4, 217-246. DOI= http://dx.doi.org/10.2753/MIS0742-1222230410

[28] Xiao, B. and Benbasat, I. 2007. E-commerce Product Recommendation Agents: Use, Characteristics, and Impact. *Mis Quarterly*. 31, 1, 137-209.