Journal of Universal Computer Science, vol. 23, no. 2 (2017), 187-207
submitted: 30/5/16, accepted: 4/1/17, appeared: 28/2/17 © *J.UCS*

Does the Users' Tendency to Seek Information Affect Recommender Systems' Performance?

Umberto Panniello
(Politecnico di Bari, Bari, Italy
umberto.panniello@poliba.it)

Lorenzo Ardito
(Politecnico di Bari, Bari, Italy
lorenzo.ardito@poliba.it)

Antonio Messeni Petruzzelli
(Politecnico di Bari, Bari, Italy
antonio.messenipetruzzelli@poliba.it)

Abstract: Much work has been done on developing recommender system (RS) algorithms, on comparing them using business metrics (such as customers' trust or perception of recommendations' novelty) and on exploring users' reactions to recommendations. It was demonstrated that different recommender systems perform differently on several performance metrics and that different users react differently to the same kind of recommendations. As a consequence, some scholars challenged to explore how users with different tendency to seek information during their purchasing process may react to different kind of recommendations. To the best of our knowledge, none of the prior works studied if users' tendency to seek information has an effect on recommender systems' performance. Different users may traditionally have different propensity to seek information and to receive suggestions and therefore they may react differently to the same recommendations. To this aim, we performed a live experiment with real customers coming from a European firm.

Keywords: Recommender system; context-aware; experiment; trust; accuracy; novelty.
Categories: H.3.3, H.3.5, L.3.2

1 Introduction

Recommender systems (RS) are traditionally used to deliver personalized product recommendations to users. They were widely studied in academia during past years. Much research has been done on proposing new recommendation algorithms and on comparing them in order to find the best performing one [Ricci et al. (2011)]. In particular, among several different recommendation engines, it was demonstrated that context-aware recommender systems significantly improve recommendations performance [Adomavicius and Tuzhilin (2011)]. Context-aware recommender systems take into account the context in which the transaction takes place [Adomavicius et al. (2005)] and they use this information to refine the generated recommendations.

In general, all these algorithms were designed to improve their predictive accuracy since the primary goal of these systems is to be as much accurate as possible in retrieving interesting items for the users [Adomavicius and Tuzhilin (2005)]. Many studies explored how users react to recommendations [Pu et al. (2012)] and it was demonstrated that different users react differently to recommendations [Komiak and Benbasat (2006)]. In particular, some studies challenged scholars to study how the tendency in seeking for information during the customer decision process may influence the recommendations performance [Knotzer and Madlberger (2007)].

During last years, recommender systems were largely adopted by online retailers, such as Netflix and Amazon. This phenomenon put the light on their business performance. In fact, RS have the potential to help firms increase their profits by converting browsers into buyers, increasing customer loyalty and company sales [Schafer et al. (2001)]. RS were born as information retrieval tools, but they have quickly become interesting for business purposes [Gorgoglione et al. (2011), Panniello et al. (2016)]. Therefore RS started to be studied and compared using customer behaviour metrics such as customers' trust, perception of recommendations' diversity, novelty [Pathak and Patra (2015)] and purchasing behaviour metrics [Gorgoglione et al. (2011)]. Many of these studies demonstrated that the comparison among RS depend on the performance metrics used to perform the comparison [Panniello et al. (2014)]. One RS may outperform another in terms of one performance metric, while it can be outperformed in terms of another metric [Wang et al. (2014)].

Therefore, there is a large knowledge of how users react to recommendations and on the effects of RS on business performance. In particular, it was demonstrated that different recommender systems perform differently on these performance metrics and that different users might react differently to the same recommendations. In addition, some scholars stated that it would be important to investigate how the customers' tendency in seeking for information during the decision process may influence the recommendations' performance. However, to the best of our knowledge, none of these works studied whether users' tendency to seek information through recommendations has an effect on recommender systems' performance. Some recommendation engines could be better suited for users with a scarce propensity to seek for suggestions while other engines could be better suited for users with a high propensity to seek for suggestions.

We aim at filling this gap. We performed a live experiment to demonstrate that RS performance can change depending on the customers' tendency to seek information. In particular, we found that the customers' tendency to seek information through recommendations has an effect on recommender systems' performance and in particular, we found that personalized recommendations outperformed random ones in terms of accuracy only when considering users that use to discuss and receive recommendations, while it was not true when considering users that decide autonomously their purchases. In addition, we found that CARS and random recommendations are perceived as more novel and generated higher level of trust with respect to content-based recommendations only for the users that use to decide autonomously their purchases. The practical and immediate implication of our research on business is that the best RS can change depending not only on the performance metric that a company wants to focus on but also on the propensity to seek information of the target group of customers.

The rest of the paper is organized as follow. First of all, we discuss the main works done during previous years on topics related to our research issue. We present the methodology followed during our experiment and then we discuss the main results obtained with our analyses.

2 Literature Background

Recommender systems are traditionally used to deliver personalized product recommendations to users. During past years many recommendation engines were proposed and compared among them in terms of their predictive accuracy [Aamir and Bhusry (2015), Ricci et al. (2011)]. This was done since the main recommender systems' task is to help users retrieving items. It was demonstrated that recommender systems reduce the total number of products examined by customers [Moore and Punji (2001)] and the number of products about which detailed information are obtained [Haubl and Trifts (2000)]. In particular, it was largely demonstrated that context-aware recommender systems (CARS) significantly outperforms the traditional recommendation engines [Adomavicius and Tuzhilin (2011)]. These RS take into account additional contextual information about the transactions performed by customers, such as time, location or intent of the transaction [Adomavicius et al. (2005)].

In this field much work has been done on studying users' reactions to recommendations [Pu et al. (2012)]. It has been demonstrated that RS users spend less time searching for information and completing the shopping task [Hostler et al. (2005), Pedersen (2000), Vijayasarathy and Jones (2001)] and they have longer actual and perceived decision time [Olson and Widing (2002)]. In addition, some studies demonstrated that the use of RS improves consumers' decision quality, in terms of preference matching scores [Hostler et al. (2005), Pereira (2001)], confidence in decision [Olson and Widing (2002), Pereira (2001)], choice of non-dominated alternatives [Haubl and Trifts (2000), Haubl and Murray (2006)] and product switching [Haubl and Trifts (2000), Haubl and Murray (2006), Olson and Widing (2002)]. In particular [Davis (1989)], proposed the technology acceptance model (TAM) to predict and explain users' intention to accept technology. After this work, many researchers applied TAM in their studies of technology adoption, which resulted in the development of several theoretical models that explain the individual intention to use technology [Venkatesh and Davis (2000)]. Starting from these models, [Venkatesh et al. (2003)] proposed the unified theory of acceptance and use of technology (UTAUT).

After these studies, some scholars have started to examine the underlying factors that influence customers' intention to adopt RS [Benbasat and Wang (2005), Cho and Sagynov (2015), Jiang et al. (2015), Sun et al. (2015)] and some evidences that different users react differently to the same recommendations were found [Ekstrand et al. (2015), Elkahky et al. (2015)]. In particular, it was found that situational or personal aspects such as product expertise [Kamis and Davern (2004), Knijnenburg and Willemsen (2009), Knijnenburg and Willemsen (2010), Knijnenburg et al. (2011)] and privacy concerns [Komiak and Benbasat (2006), Teltzrow and Kobsa (2004)] can influence how people interact with and evaluate the RS. Product category knowledge is negatively related to perceived ease of use and perceived usefulness of the decision tools [Kamis and Davern (2004)]. For users with low product knowledge, a needs-based RS results in higher trust [Komiak and Benbasat (2006)]. Users with high product class knowledge have more positive affective reactions (trust and satisfaction) to the content-filtering RS than the collaborative-filtering ones, the reverse is true for users with low product class knowledge [Pereira (2001)]. Even if much work has been done, several aspects still need to be investigated. In particular, [Knotzer and Madlberger (2007)] demonstrated that opinion seeking tendency positively influences a consumer's interest in reading product details and they challenged to learn more about how information seekers differ from non-seekers.

During the last years RS were adopted by many big e-commerce players, such as Netflix and Amazon [Gomez-Uribe and Hunt (2015)]. As a consequence, many studies compared existing recommender systems in terms of business-based metrics instead of accuracy-based ones. Several scholars have acknowledged that prediction accuracy alone is not sufficient to fulfil user satisfaction, build user loyalty, or persuade customers to purchase products [Benbasat and Wang (2005), Xiao and Benbasat (2007)]. Therefore, there is a pressing need to understand factors beyond prediction accuracy that influence the user experience (i.e. users' subjective evaluation toward using recommender systems) and the effectiveness of recommender system use [Konstan and Riedl (2012), Swearingen and Sinha (2001), Xiao and Benbasat (2007)]. There has been much research on this perspective since [Schafer et al. (1999)] challenged scholars to design an RS to both maximize customer utility and business value at the same time. Some scholars studied what factors affect an RS adoption by customers [Cooke et al. (2002), Liang et al. (2007)]. Taking a few steps ahead along this direction, some scholars have studied how RS affect certain aspects of customers' behaviour even closer to business goals, such as intentions and attitudes. For instance, [Pu et al. (2011)] found that ease of use and perceived usefulness is important for usage intentions, while trust and choice confidence are crucial for purchasing intentions. [Bharati and Chaudhury (2004)] studied how relevance, accuracy, completeness, and timeliness of recommendations have a significant effect on users' decision-making and satisfaction. Other scholars have directly investigated the economic effect of RS usage on the business. [Schafer et al. (2001)] argued that recommender systems help increase sales by converting browsers into buyers, increasing cross-selling opportunities, and building customer loyalty. [Fleder and Hosanagar (2009)] demonstrated that RS that discount item popularity in the selection of recommendable items may increase sales more than RS that do not. [Pathak et al. (2010)] found that the strength of recommendations has a positive impact on sales. [Gorgoglione et al. (2011)] showed the effect of recommendations on customers' purchasing behaviour and business sales. Finally, many of the works in this field demonstrated that the comparison among RS depend on the performance metrics used to perform the comparison [Liang et al. (2007), Panniello et al. (2014), Wang et al. (2014)].

As a result, there is large knowledge of how users react to recommendations and on the effects of RS on business performance. It was largely demonstrated that different recommender systems perform differently on these performance metrics and it was also demonstrated that different users react differently to RS. In particular, some scholars highlighted that it would be very important to investigate how the tendency in seeking for information during the customer decision process may influence the recommendations performance [Knotzer and Madlberger (2007)]. However, to the best of our knowledge, none of these works studied whether comparisons across different recommendation engines vary when considering users with different tendency to seek information through recommendations. In other words, none of these works studied whether users' intent to seek information through recommendations has an effect on recommender systems' performance. On one side, users with different tendency to seek information through recommendations may react differently to the same recommendations. Some users may be more disposed to seek suggestions while other may be less thus reacting differently to the same recommendations. On the other side, as largely demonstrated by previous literature, different recommendations lead to different performance. In our case, users with the same tendency to seek information through recommendations may react differently to different recommendations. Users more disposed to seek suggestions may prefer some type of recommendations while users less disposed to seek suggestions may prefer different type of recommendations. Therefore, our research aim at filling this gap: we want to study whether different users' tendency to seek information through suggestions lead to different performance and we want to study this point by varying the type of recommendation engine.

3 Methodology

In order to answer our research issue, i.e. exploring whether different users' tendency to seek information through suggestions affect recommendations' performance, we conducted an experiment in a real-world setting in partnership with a well-known European firm operating in the publishing industry worldwide. The company's Web division mainly sells comic books and related products, such as DVDs, stickers, and T-shirts. As a part of its normal business, the company sends a weekly non-personalized newsletter to approximately 23,000 customers and agreed to send personalized recommendations of comic books via e-mail to a sample of this customer base as a part of our project. Since we want to explore our research issue varying also the type of recommendation engine used to generate recommendations, during the experiment we used three different RS, a content-based, a context-aware and a random recommender namely. We have chosen a content-based recommendation algorithm, rather than a collaborative filtering (CF) method, because it would have been difficult to generate meaningful recommendations using the CF approach since the experiment was carried out with few participants and the user/item matrix was relatively sparse – the two conditions adversely affecting CF results.

The content-based recommendation engine basically recommends items that are similar to the ones the user preferred in the past. As defined in literature [Pazzani and Billsus (2007)], it computes rating $r(u,s)$ of item s for user u based on the ratings $r(u,s_j)$ assigned by user u to items $s_j \in S$ that are similar to item s. In particular, let *ItemProfile(s)* for item s and *UserProfile(u)* for user u, be two vectors representing the item characteristics and the customer preferences, respectively. *ItemProfile(s)* is computed by extracting a set of keywords from a description of item s. The keywords describe the item and its contents, including author and publisher details. *UserProfile(u)* is computed by analysing the content of the items previously seen and rated by user *u*. In particular, the vector is defined as a vector of weights (w_{u1}, …, w_{uj} , …, w_{uz}), where each w_{uj} denotes the importance of keyword *j* to user *u*. We computed w_{uj} as an "average" of the ratings provided by user *u* to those items that contained the keyword $j \in Z$. Candidate items are compared with user profile and the most similar items are recommended. We compute relevance *r(u,s)* of item *s* to user *u* by matching the *UserProfile(u)* and the *ItemProfile(s)*. The top 10 items with the highest score are presented (recommended) to the user in the newsletter.

Since our aim was to fairly compare a traditional (content-based) RS with a CARS, the CARS developed for our experiment used the same content-based algorithm discussed above. The only difference is that we used *UserProfile(u,k)* which is the profile of user *u* in context *k* instead of *UserProfile(u)* which does not consider the context *k*. We computed profile *UserProfile(u,k)* using the pre-filtering approach [Adomavicius and Tuzhilin (2011), Panniello et al. (2009)] by analysing the content of the items previously seen and rated by user *u* in context *k*. In particular, the contextual information *k* is used as a label for filtering out those items that were not rated in this context *k*, i.e., this method selects from the initial set of all the ratings only those referring to context *k*. As a result, *UserProfile(u,k)* contains only the data pertaining to context *k*. After that, the content-based algorithm is launched only on these selected data to produce recommendations specific to context *k*. We follow [Panniello et al. (2009)] and [Adomavicius and Tuzhilin (2011)] by defining context with a set of contextual attributes (variables) as follows. First, we assume that domain of contextual attribute *K* is defined by a set of *q* attributes $K = (K_1,\ldots, K_q)$ having a hierarchical structure associated with it. The values taken by attribute K_q define finer levels, while K_1 coarser levels of contextual knowledge [Kwon and Kim (2009)]. In our experiment, we used two distinct contextual variables: the "intent of a purchase" made by a customer and the "customer's mood". We set "intent of purchase" and "mood" as contextual variables in our study after setting up focus groups, conducting several interviews with readers of comic books, and discussing the produced results with the company management. In particular, most of the interviewees told us that they modify their behaviour depending on the intent of purchase and that their choice of reading a certain comic book is related to the emotional content and may depend on their mood (i.e., since customer's behaviour and preferences change depending on these variables, we can consider them as contextual variables [Adomavicius et al. (2005)]). In addition, we have found similarity between these findings and several web sites settings, such as the "wish list" and the "gift options" of certain e-malls, and the "mood menu" of several music vendors and providers (e.g., see the LastFM and Musicovery examples). All this supports our choice of contextual variables "intent of purchase" and "mood" in the experiments and their importance in other RS. These two variables are presented in Figure 1. The "intent of purchase" contextual variable distinguishes whether the user is looking for recommendations for his/her personal interest (further distinguishing between recommendations for his/her collections, special issues or occasional readings) or for a gift (further distinguishing between recommendations for a gift to a partner, a friend, etc.).

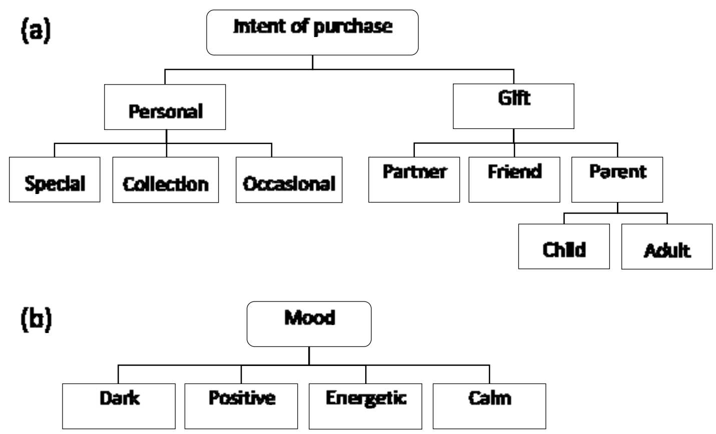

Figure 1: Hierarchical structure of contextual variables (a) intent of purchase and (b) customers' mood.

Contextual variable "mood" distinguishes between different moods in which the customer can be. In particular, the customer may like to receive recommendations depending on his/her current mood that can be dark, energetic, positive or calm in our study. When users of the contextual treatment group received the newsletter, it was asked them to specify the context in which they were. In particular, users had the possibility to choose if they wanted to receive recommendations for a specific purchasing intent or for a specific mood in which they were. In the first case, users had the possibility to ask recommendations for a personal intent or for a gift (then detailing their choice). In the second case, users had the possibility to ask recommendations for a specific mood in which they were. Then, only recommendations for the specified context were shown to the participants. We decided to explicitly ask users to declare the context, since our aim was to evaluate the quality of the contextual recommender system in terms of recommending items for a specific context. In other words, asking the user to declare the context for which he/she wants to receive recommendations makes us sure that he/she will evaluate the recommendations in terms of how suitable are them for the declared context. If we do not do so, the customer could evaluate each specific recommendation in general thus not giving us the opportunity to measure whether the specific recommendation is good or not for the declared context. In other words, if we do not ask the context, the recommender systems would work as a traditional one instead of a CARS. In addition, if we do not explicitly ask users to declare the context, it would not be possible to associate the obtained feedbacks to the corresponding context.

It was possible to show recommendations pertaining to the chosen context since recommendations were generated for each user and each context but only those pertaining to the context selected by the user were showed. Therefore, it was also possible to change the target context once it was set and customers could see and rate the recommended items also in another context.

Unlike the content-based and context-aware approaches, the random approach does not take the user profile into consideration when recommending new products. Instead, it randomly selects a set of items from those that have not been recommended or purchased before.

The 360 customers were randomly selected from the customer database. According to the privacy laws, the firm asked customers to state explicitly if they wanted to join this project to improve the customer service. The activity was presented as collaboration between the company and a university aimed at improving the customer service. The experiment participants were then randomized into three experimental treatment groups, each of the three groups receiving either random, content-based, or context-aware recommendations. We performed statistical tests (t-test and chi-square test) on the age and gender variables among the treatment groups and between them and the whole population, and we found no statistically significant differences. Therefore we concluded that the sample selection was unbiased. Each subject received a personalized weekly newsletter displaying 10 recommended comic books for nine consecutive weeks. It contained a link to a personal recommendation page displaying the ten recommended items. Each item was presented with the following information: title, cover image, description and a "see more details" link. The customers were invited to rate each recommended product by clicking on a (0-5) point scale. These ratings were used to update the user profile for each user.

Since our aim is to measure whether different types of recommender systems perform differently based on how users seek information during their purchasing process, we measured a set of information useful to explore this research issue. In particular, we asked users how they traditionally use to make their comic books purchase decisions. Each user had the possibility to declare that "I decide completely on my own" or "I prefer to discuss with someone (friends or other) and listen their recommendations". This question was used to label users based on how they seek information during their purchasing process, i.e. "autonomous" and "discussing" namely. This variable was not used as a contextual variable. Even if the purchasing decision process can be more complex than deciding to receive or not suggestions, we adopted this basic distinction in order to demonstrate whether the presence or absence of tendency to seek information through suggestions has an effect on the recommendations' performance. Investigating what are the effects of different levels of tendency to seek information on recommendations' performance is beyond the aim of this research and it would require different experiments and analyses.

In order to compare RS on different performances, we decided to measure some of the most popular metrics in recent literature on recommender systems. These metrics are recommendations' accuracy, recommendations' perceived novelty and users' trust on the recommender system.

Accuracy was measured by precision and average ratings. Among the traditional information retrieval performance metrics, such as precision, recall and F-measure, only precision could be computed in our case, since it was not possible to know the ratings of the unseen items needed to compute the recall and the F-measure. According to [Herlocker et al. (2004)], precision was measured as:

$$P = \frac{N_{rs}}{N_s}$$

where N_s is the total number of items recommended to the customer (selected by the RS as items to be recommended) and N_{rs} is the number of items which proved to be "relevant" (good recommendations) for the customer among those selected by the RS. We considered an item being "relevant" if it was rated as three, four or five on the 0–5 rating scale used during our experiment. We decided to consider items rated as 3, 4 or 5 as relevant instead of considering only items rated as 4 or 5 as discussed in [Herlocker et al. (2004)] since our rating scale was from 0 to 5 instead of 1-5 scale (as in [Herlocker et al. (2004)]). As previously mentioned, we also measured accuracy as the average rating *(*$Averagerating_{z,u}$*)* provided by user u in period z over the rated items. In particular, being $rating_{s,z,u}$ the rating provided by user u to item s in period z and S the total number of items rated by u in period z, the average rating for user u in period z is measured as:

$$AverageRating_{z,u} = \frac{\sum_S rating_{s,z,u}}{S}$$

Recommendations' perceived novelty was measured by explicitly asking to the customers whether "personalized newsletters recommended comic books that the user didn't know". This question is part of the final survey provided to the participants of the experiment for measuring how much the participants trusted the received recommendations (see Table 1).

	Questions in the survey
Q_1	I usually trust people
Q_2	This personalized newsletter is like a real expert in assessing comic books
Q_3	Personalized newsletters provided me with relevant recommendations
Q_4	Personalized newsletters recommended comic books that I didn't know
Q_5	I am willing to let this newsletter assist me in deciding which product to buy
Q_6	The newsletter is reliable
Q_7	I trust the personalized newsletter
Q_8	The company created the personalized newsletter to help me
Q_9	The personalized newsletter is a service provided by the company to customers
Q_{10}	I bought some of the recommended products offline
Q_{11}	I think the recommended products were expensive

Table 1: Final survey

The questions were composed according to the literature on experimental design and on trust [Beldad et al. (2010), Benbasat and Wang (2005), Doney et al. (1998), Mayer et al. (1995), Schoorman et al. (2007)]. Each answer was provided in the (1-5) scale. In particular, questions Q_1, Q_{10} and Q_{11} were used to check that there were not any biases, in terms of general propensity to trust and off-line purchases, into the data. In fact, different levels of propensity to trust and different amount of off-line purchases may affect the other results.

Even if it would be very interesting to compare users and RS also in terms of generated sales, it was not possible since the experiment lasted only for nine weeks thus resulting in few sales data points that are too small for analyses on different groups of customers (i.e., autonomous and discussing) and RS (i.e., content-based, CARS and random).

From the whole customer base obtained after the experiment, we selected only the users for whom we have full information (i.e., ratings, response to final survey and information about their purchasing process) thus resulting in 143 users.

4 Results

As stated in previous sections, we know from prior literature that different recommender systems perform differently on specific performance metrics and that different users react differently to the same recommendations. In addition, prior research challenged scholars to investigate how tendency in seeking for information during the decision process may influence the recommendations performance. Therefore, our aim is to compare different recommender systems on different performance metrics when considering users with different tendency in seeking for information during their purchasing process. Therefore, this section is organized as follow: for each performance metric, we present the results obtained from our experiment and the insights obtained when splitting customers based on how they seek information during their purchasing process.

Regarding the recommendations' accuracy, in general, personalized recommendations (content and contextual ones) are more accurate than random ones [Chen and Pu (2010), Knijnenburg and Willemsen (2010), Ziegler et al. (2005)] because users are aware of the quality distinctions between random and personalized recommendations [Knijnenburg et al. (2012)]. Our results confirm prior research: precision and average ratings of users who received personalized recommendations are significantly higher than those of users receiving random recommendations (see Table 2).

	# of observations	Avg. Ratings (standard deviation)	Avg. Precision (standard deviation)
Content-based	44	3.09 (.675)	.58 (.197)
CARS	65	3.04 (.730)	.54 (.198)
Random	34	2.64 (.772)	.40 (.200)
	Difference in mean avg. ratings (t-value)	Difference in mean avg. Precision (t-value)	
Content-based vs. CARS	.051 (.375)	.038 (1.066)	
Content-based vs. Random	.450 (2.742)**	.178 (3.919)***	
CARS vs. Random	.398 (2.527)*	.136 (3.245)**	

Table 2: Average ratings and precision of the three recommendation engines

In particular, as Table 2 demonstrates, the average ratings and precision generated by the personalized recommendations (content-based and context-aware) are higher than the random ones. This is confirmed by the fact that the differences between content-based and random groups and between the context-aware and the random groups were found to be statistically significant. On the contrary, there are no statistically significant differences between the two groups receiving personalized recommendations. In order to explore our research issue, we split customers based on how they seek information during their purchasing process (using the corresponding question, see previous section) and we measured and statistically validated differences between these groups of users.

	# of observations	Avg. Ratings (standard deviation)	Avg. Precision (standard deviation)
Content-based	32	3.03 (.731)	.56 (.205)
CARS	30	2.87 (.726)	.50 (.179)
Random	19	2.76 (.899)	.40 (.167)
	Difference in mean avg. ratings (t-value)	Difference in mean avg. Precision (t-value)	
Content-based vs. CARS	.159 (.860)	.066 (1.345)	
Content-based vs. Random	.267 (1,160)	.164 (2.952)**	
CARS vs. Random	.108 (.465)	.098 (1.913)	

Table 3: Average ratings and precision of the three recommendation engines when considering "autonomous" users

	# of observations	Avg. Ratings (standard deviation)	Avg. Precision (standard deviation)
Content-based	12	3.27 (.478)	.63 (.174)
CARS	35	3.19 (.711)	.58 (.208)
Random	15	2.50 (.567)	.41 (.243)
	Difference in mean avg. ratings (t-value)	Difference in mean avg. Precision (t-value)	
Content-based vs. CARS	.078 (.354)	.055 (.824)	
Content-based vs. Random	.772 (3.761)**	.223 (2.675)*	
CARS vs. Random	.694 (3.346)**	.168 (2.482)*	

Table 4: Average ratings and precision of the three recommendation engines when considering "discussing" users

We found that personalized recommendations are more accurate than random ones only for users that use to discuss and to accept recommendations during their purchasing process (see Table 4). In particular, both the average ratings and precision are statistically higher for the personalized recommendations than for the random ones when considering this set of users (see Table 4). On the contrary, both average ratings and precision are not statistically different between personalized and random recommendations when considering users that use to decide autonomously, with the exception of the average precision between content and random groups (see Table 3). From a managerial point of view, this is an important result since it confirms that it is useful to use a personalized recommender system only when customers tend to seek information through a discussion. On the contrary, if the users use to decide autonomously without seeking information through suggestions, it results in no differences in terms of accuracy across personalized and un-personalized recommendations.

Regarding perceived recommendations' novelty, previous research demonstrated that contextual and random recommendations are perceived as more novel than content ones because of their diversity [Panniello et al. (2014)]. In fact, even if the random recommendations are less accurate than personalized ones (i.e., content or contextual), these are perceived as more novel because are randomly generated. Our results confirm prior research since context and random recommendations were perceived as more novel than content ones (see Table 5).

	#of observations	Avg. Perceived novelty (standard deviation)		Difference in mean avg. perceived novelty (t-value)
Content-based	44	2.61 (1.385)	Content vs. CARS	.448 (1.694)*
CARS	65	3.06 (1.333)	Content vs. Random	.769 (2.377)*
Random	34	3.38 (1.457)	CARS vs. Random	.321 (1.101)

Table 5: Perceived novelty of the three recommendation engines

As we have done with the accuracy metrics, we split customers based on how they seek information during their purchasing process and we measured and statistically validated differences between these groups of users.

	#of observations	Avg. Perceived novelty (standard deviation)		Difference in mean avg. perceived novelty (t-value)
Content-based	32	2.63 (1,314)	Content vs. CARS	.775 (2.354)*
CARS	30	3.40 (1.276)	Content vs. Random	.849 (2.070)*
Random	19	3.47 (1.577)	CARS vs. Random	.074 (.180)

Table 6: Perceived novelty of the three recommendation engines when considering "autonomous" users

We found that contextual and random recommendations were perceived as more novel than content ones only by users who decide autonomously while it was not confirmed for the others. This finding can be explained by the fact that in some cases users do not necessarily perceive the diversified recommendations as more novel [Bollen et al. (2010), Ziegler et al. (2005)]. This finding is interesting since it means that perceived recommendations' novelty varies depending on how users seek information. In particular, users that use to decide autonomously their purchases perceive context-aware and random recommendations as more novel than content-based ones. On the contrary, users that use to seek information through discussions and recommendations do not perceive any difference in terms of novelty across the recommendations engines. It has been largely demonstrated that the perceived recommendations' novelty is an important driver for business performance and that some kind of recommender systems (i.e., CARS) outperforms others in terms of this metric [Panniello et al. (2014)]. Our results prove that it depends on the users' tendency to seek information through recommendations.

	# of observations	Avg. Perceived novelty (standard deviation)		Difference in mean avg. perceived novelty (t-value)
Content-based	12	2.58 (1.621)	Content vs. CARS	.188 (.400)
CARS	35	2.77 (1.330)	Content vs. Random	.683 (1.202)
Random	15	3.27 (1.335)	CARS vs. Random	.495 (1.205)

Table 7: Perceived novelty of the three recommendation engines when considering "discussing" users

Regarding customers' trust, previous research demonstrated that users perceive some differences among different types of recommender systems since it is mediated by other factors, such as accuracy or novelty [Knijnenburg et al. (2011)]. The results of our final survey confirm previous research (see Table 8).

	Content-based Avg. (standard deviation)	CARS Avg. (standard deviation)	Random Avg. (standard deviation)
Q1	2.80 (1.069)	3.08 (.989)	2.94 (.814)
Q2	3.05 (.806)	3.05 (1.124)	2.97 (1.159)
Q3	3.41 (.816)	3.51 (1.033)	3.44 (1.133)
Q5	2.59 (1.064)	3.09 (1.271)	3.15 (1.395)
Q6	3.05 (.939)	3.35 (1.052)	3.53 (1.107)
Q7	2.86 (.915)	3.20 (1.107)	3.44 (1.236)
Q8	3.48 (1.210)	3.71 (1.114)	3.45 (1.277)
Q9	4.07 (.985)	4.02 (1.008)	3.88 (1.053)
Q10	3.82 (1.435)	3.65 (1.515)	3.24 (1.542)
Q11	2.86 (1.503)	2.86 (1.540)	2.97 (1.468)
	Content vs. CARS Difference in mean avg. (t-value)	Content vs. Random Difference in mean avg. (t-value)	CARS vs. Random Difference in mean avg. (t-value)
Q1	.281 (1.411)	.146 (.660)	.136 (.687)
Q2	.001 (.004)	.076 (.338)	.076 (.315)
Q3	.099 (.531)	.032 (.145)	.067 (.294)
Q5	.501 (2.154)*	.561 (2.002)*	.059 (.211)
Q6	.308 (1.567)	.484 (2.087)*	.176 (.774)
Q7	.340 (1.669)*	.581 (2.369)*	.241 (.989)
Q8	.230 (1.023)	.023 (.080)	.253 (1.011)
Q9	.054 (.277)	.191 (.813)	.137 (.625)
Q10	.172 (.594)	.576 (1.688)	.404 (1.240)
Q11	.002 (.007)	.106 (.309)	.108 (.334)

Table 8: Trust of the three recommendation engines

Answers to Q1, Q10 and Q11 are not statistically different among groups thus reassuring that no biases in terms of general propensity to trust and off-line purchases may have affected other results. The other results suggest that the content-based recommendations generate lower level of trust in comparison to CARS and random ones (see Q6 and Q7 in Table 8). This result confirms prior works suggesting that customers' trust is mediated by several factors such as accuracy and novelty [Knijnenburg et al. (2011)]. In addition, users receiving CARS and random recommendations are more disposed to let the newsletter assists them in deciding which products to buy (see Q5 in Table 8) than users receiving content-based recommendations.

As done previously, we split customers based on how they seek information during their purchasing process and we measured and statistically validated differences between these groups of users.

	Content-based Avg. (standard deviation)	CARS Avg. (standard deviation)	Random Avg. (standard deviation)
Q1	2.69 (1.148)	3.13 (1.137)	3.00 (.816)
Q2	3.03 (.861)	3.07 (1.172)	3.21 (1.084)
Q3	3.50 (.718)	3.63 (.964)	3.63 (1.116)
Q5	2.56 (1.014)	3.13 (1.224)	3.11 (1.568)
Q6	3.06 (.878)	3.43 (.898)	3.63 (1.165)
Q7	2.87 (.846)	3.20 (1.031)	3.53 (1.307)
Q8	3.47 (1.164)	3.73 (.944)	3.63 (1.116)
Q9	4.19 (.821)	4.13 (.819)	4.00 (1.000)
Q10	3.78 (1.475)	3.50 (1.548)	3.21 (1.619)
Q11	2.84 (1.526)	2.63 (1.608)	2.79 (1.512)
	Content vs. CARS Difference in mean avg. (t-value)	Content vs. Random Difference in mean avg. (t-value)	CARS vs. Random Difference in mean avg. (t-value)
Q1	.446 (1.535)	.313 (1.039)	.133 (.443)
Q2	.035 (.136)	.179 (.652)	.144 (.431)
Q3	.133 (.620)	.132 (.513)	.002 (.006)
Q5	.571 (2.005)*	.549 (1.503)	.002 (.055)
Q6	.371 (1.644)	.569 (1.979)*	.198 (.671)
Q7	.329 (1.365)	.655 (2.156)*	.326 (.973)
Q8	.265 (.979)	.163 (.490)	.102 (.342)
Q9	.054 (.260)	.188 (.727)	.133 (.509)
Q10	.281 (.732	.571 (1.288)	.289 (.627)
Q11	.210 (.529)	.054 (.123)	.156 (.339)

Table 9: Trust of the three recommendation engines when considering "autonomous" users

Splitting customers, we found that users that decide autonomously (see Table 9) trusted differently the different recommender systems, while users that seek information through discussion and recommendations did not trust differently the different recommender systems (see Table 10). In other words, users with different tendency to seek information trusted differently the systems. In particular, users that use to discuss before purchases did not show different levels of trust among different recommendation engines (Table 10), while users that use to decide autonomously demonstrated a higher propensity to let CARS assists them in comparison to random and content recommendations, and they trusted more the random recommendations (Table 9).

	Content-based Avg. (standard deviation)	CARS Avg. (standard deviation)	Random Avg. (standard deviation)
Q1	3.08 (.793)	3.03 (.857)	2.87 (.834)
Q2	3.08 (.669)	3.03 (1.098)	2.64 (1.216)
Q3	3.17 (1.030)	3.40 (1.090)	3.20 (1.146)
Q5	2.67 (1.231)	3.06 (1.327)	3.20 (1.207)
Q6	3.00 (1.128)	3.29 (1.178)	3.40 (1.056)
Q7	2.83 (1.115)	3.20 (1.183)	3.33 (1.175)
Q8	3.50 (1.382)	3.69 (1.255)	3.21 (1.477)
Q9	3.73 (1.348)	3.91 (1.147)	3.71 (1.139)
Q10	3.92 (1.379)	3.77 (1.497)	3.29 (1.490)
Q11	2.92 (1.505)	3.06 (1.474)	3.21 (1.424)
	Content vs. CARS Difference in mean avg. (t-value)	Content vs. Random Difference in mean avg. (t-value)	CARS vs. Random Difference in mean avg. (t-value)
Q1	.055 (.194)	.217 (.686)	.162 (.617)
Q2	.055 (.162)	.440 (1.117)	.386 (1.078)
Q3	.233 (.648)	.033 (.078)	.200 (.586)
Q5	.390 (.895)	.533 (1.131)	.143 (.358)
Q6	.286 (.733)	.400 (.949)	.114 (.324)
Q7	.367 (.939)	.500 (1.124)	.133 (.366)
Q8	.186 (.431)	.286 (.506)	.471 (1.129)
Q9	.187 (.452)	.013 (.026)	.200 (.552)
Q10	.145 (.296)	.631 (1.114)	.486 (1.027)
Q11	.140 (.283)	.298 (.518)	.157 (.340)

Table 10: Trust of the three recommendation engines when considering "discussing" users

In conclusion, we found that the users' tendency to seek information during their purchasing process has an effect on recommender systems' performance. In fact, we found that personalized recommendations are perceived as more accurate in comparison to random ones only for users that prefer to discuss and to accept recommendations. On the contrary, no differences in terms of accuracy between personalized and random recommendations were found for users that use to decide autonomously their purchases. We also found that CARS and random recommendations are perceived as more novel than content-based ones only by the users that decide autonomously their purchases, while it is not the case for the users that use to discuss and to accept suggestions. Finally, we have found that users that decide autonomously their purchases demonstrated a higher trust for CARS and random recommendations than for content-based ones, while users that use to discuss and to accept suggestions demonstrated similar levels of trust for all the recommendation engines.

It would be interesting to investigate the reasons beyond these findings but it would require additional and different experiments that are far from the aim of this study. In fact, we aimed at demonstrating that differences in recommendations' performance do exist when considering users with different tendency to seek information, while we did not aim at finding the drivers of these differences. Investigating this point would require additional and different experiments and it is a very interesting research issue for further research.

All these findings can be easily used in real world recommender systems since it was previously demonstrated that personality-based recommenders are more likely to be accepted from users and that they reduce users' perceived cognitive effort and increase their satisfaction [Hu and Pu (2009), Hu and Pu (2009)].

5 Conclusions

Prior literature on recommender systems has demonstrated that different recommender systems perform differently on specific performance metrics and that different users react differently to the same recommendations. In particular, among the users' characteristics used to explore this issue, some scholars stated that it would be interesting to study how the tendency in seeking for information during the customer purchasing process may influence the recommendations' performance. To the best of our knowledge, none of the previous studies explored this issue. Therefore, our aim was to study the aforementioned phenomenon when considering users with different tendency in seeking information during their purchasing process.

We conducted a live experiment with real customers of a European e-commerce website. In particular, we had three different groups of customers receiving three different types of recommendations (namely content-based, context-aware and random recommendations). We asked to the customers of these three groups how they usually seek for information during their purchasing process. They may decide completely on their own or they may discuss with someone and listen their recommendations. We sent recommendations to them for nine weeks and we measured several metrics. In particular, we measured recommendations' accuracy, perceived recommendations' novelty and users' trust in the recommender system. Finally, we split customers based on their tendency in seeking information and we compared the three recommendation engines in terms of the aforementioned metrics.

We found that the tendency to seek information has an effect on recommender systems' performance. In fact, we found that personalized recommendations outperformed random ones in terms of accuracy only when considering users that use to discuss and receive recommendations, while it was not true when considering users that decide autonomously their purchases. In addition, we found that CARS and random recommendations are perceived as more novel and generated higher level of trust with respect to content-based recommendations only for the users that use to decide autonomously their purchases. All these results were obtained in a specific empirical context (i.e., an e-commerce selling comic books) and therefore several characteristics of this setting cannot be generalized to other settings (such as, the chosen contextual variables). However, since our aim was to demonstrate that different approaches in seeking information result in different recommender systems' performance while it was not to define the characteristics under which it is true, we maintain that our results are not affected by generalizability issue. However, as further research, it would be useful to replicate our analyses using different samples and websites in order to study whether additional insights do exist.

All these results have practical and immediate implications on both business and academia since they clearly demonstrate that the choice of the RS algorithm depends not only on the performance metric on which we want to focus on but also on the tendency to seek information of the target group of customers. From a managerial point of view, we demonstrated that it is useful to use a personalized recommender system to improve the recommendations' accuracy only when customers adopt any kind of discussion while it is not true when the customers use to decide autonomously. We also demonstrated that CARS and random recommendations improve the perceived novelty only when considering users that usually decide autonomously their purchases, while it is not true when considering users that usually discuss and accept suggestions. Similar results were found in terms of customers' trust, in fact, it is useful to use CARS and random recommendations to improve customers' trust only when the users use to decide autonomously their purchases, while it is not the case when the users use to discuss and accept recommendations.

The present work aims at demonstrating that the users' tendency to seek information through recommendations has a significant effect on recommender systems' performance and it puts some lights on this phenomenon. However, it has several limitations that call for further research. First of all, other performance metrics need to be investigated. Previous research has demonstrated that recommendation engines affect many other performance metrics. We have explored the most used metrics but many others can be explicitly measured with additional experiments, such as the customer satisfaction, the customer experience or the generated sales. In addition, an extensive comparison among different recommendation engines need to be performed. During our experiment, we had the opportunity to test three recommendation engines and we selected the most used ones, but it would be interesting to extend our analysis to other engines, such as collaborative filtering one. We had the opportunity to send recommendations and survey to real customers but it was a time-consuming task resulting in few customers joining the experiment. Therefore, it would be interesting to extend the set of users involved into a live experiment as further research. In addition, it would be interesting to investigate the reasons beyond the results obtained in this work. In fact, it would be useful to conduct additional experiments to investigate the drivers of the results found in the present work. Finally, it would be interesting to study whether the findings of our analysis change over time with the customer lifetime value.

References

[Aamir and Bhusry (2015)] M. Aamir and M. Bhusry: "Recommendation System: State of the Art Approach"; International Journal of Computer Applications, 120, 12 (2015), 25 - 32.

[Adomavicius et al. (2005)] G. Adomavicius, R. Sankaranarayanan, S. Sen and A. Tuzhilin: "Incorporating Contextual Information in Recommender Systems Using a Multidimensional Approach"; ACM Transactions on Information Systems, 23, 1 (2005), 103-145.

[Adomavicius and Tuzhilin (2005)] G. Adomavicius and A. Tuzhilin: "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions"; IEEE Transactions on Knowledge and Data Engineering, 17, 6 (2005), 734-749.

[Adomavicius and Tuzhilin (2011)] G. Adomavicius and A. Tuzhilin: "Context-Aware Recommender Systems"; Recommender Systems Handbook, (2011), 217-253.

[Beldad et al. (2010)] A. Beldad, M. de Jong and M. l. Steehouder: "How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust"; Computers in Human Behavior, 26, 5 (2010), 857-869.

[Benbasat and Wang (2005)] I. Benbasat and W. Wang: "Trust in and adoption of online recommendation agents"; Journal of the Association for Information Systems, 6, 3 (2005), 72-100.

[Bharati and Chaudhury (2004)] P. Bharati and A. Chaudhury: "An empirical investigation of decision-making satisfaction in web-based decision support systems"; Decision Support Systems, 37, 2 (2004), 187-197.

[Bollen et al. (2010)] D. Bollen, B. P. Knijnenburg, M. C. Willemsen and M. Graus: "Understanding choice overload in recommender systems"; Proc. Proceedings of the fourth ACM conference on Recommender systems, Barcelona, Spain (2010), 63-70.

[Chen and Pu (2010)] L. Chen and P. Pu: "Experiments on the preference-based organization interface in recommender systems"; ACM Trans. Comput.-Hum. Interact., 17, 1 (2010), 1-33.

[Cho and Sagynov (2015)] Y. C. Cho and E. Sagynov: "Exploring Factors That Affect Usefulness, Ease Of Use, Trust, And Purchase Intention In The Online Environment"; 2015, 19, 1 (2015), 16.

[Cooke et al. (2002)] A. Cooke, H. Sujan, M. Sujan and B. Weitz: "Marketing the Unfamiliar: The Role of Context and Item-Specific Information in Electronic Agent Recommendations"; Journal of marketing research, 39, 4 (2002), 488-497.

[Davis (1989)] F. Davis: "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology"; MIS Quarterly, 13, 3 (1989), 319-340.

[Doney et al. (1998)] P. M. Doney, J. P. Cannon and M. R. Mullen: "Understanding the influence of national culture on the development of trust."; The Academy of Management Review, 23, 3 (1998), 601-620.

[Ekstrand et al. (2015)] M. D. Ekstrand, D. Kluver, F. M. Harper and J. A. Konstan: "Letting Users Choose Recommender Algorithms: An Experimental Study"; Proceedings of the 9th ACM Conference on Recommender Systems, (2015), 11-18.

[Elkahky et al. (2015)] A. M. Elkahky, Y. Song and X. He: "A Multi-View Deep Learning Approach for Cross Domain User Modeling in Recommendation Systems"; Proceedings of the 24th International Conference on World Wide Web, (2015), 278-288.

[Fleder and Hosanagar (2009)] D. Fleder and K. Hosanagar: "Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity"; Management Science, 55, 5 (2009), 697-712.

[Gomez-Uribe and Hunt (2015)] C. A. Gomez-Uribe and N. Hunt: "The Netflix Recommender System: Algorithms, Business Value, and Innovation"; ACM Trans. Manage. Inf. Syst., 6, 4 (2015), 1-19.

[Gorgoglione et al. (2011)] M. Gorgoglione, U. Panniello and A. Tuzhilin: "The effect of context-aware recommendations on customer purchasing behavior and trust"; Proc. fifth ACM conference on Recommender systems, Chicago, Illinois, USA (2011), 85-92.

[Haubl and Trifts (2000)] G. Haubl and V. Trifts: "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids"; Marketing Science, 19, 1 (2000), 4-21.

[Haubl and Murray (2006)] G. Haubl and K. B. Murray: "Double Agent: Assessing the Role of Electronic Product-Recommendation Systems"; Sloan Management Review, 47, 3 (2006), 8-12.

[Herlocker et al. (2004)] J. L. Herlocker, J. A. Konstan, L. G. Terveen and J. T. Riedl: "Evaluating collaborative filtering recommender systems"; ACM Transactions on Information Systems, 22, 1 (2004), 5-53.

[Hostler et al. (2005)] R. E. Hostler, V. Y. Yoon and T. Guimaraes: "Assessing the impact of internet agent on end users' performance"; Decision Support Systems, 41, 1 (2005), 313-323.

[Hu and Pu (2009)] R. Hu and P. Pu: "A comparative user study on rating vs. personality quiz based preference elicitation methods"; Proc. Proceedings of the 14th international conference on Intelligent user interfaces, Sanibel Island, Florida, USA (2009), 367-372.

[Hu and Pu (2009)] R. Hu and P. Pu: "Acceptance issues of personality-based recommender systems"; Proc. Proceedings of the third ACM conference on Recommender systems, New York, New York, USA (2009), 221-224.

[Jiang et al. (2015)] P. Jiang, Y. Zhu, Y. Zhang and Q. Yuan: "Life-stage Prediction for Product Recommendation in E-commerce"; Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (2015), 1879-1888.

[Kamis and Davern (2004)] A. Kamis and M. J. Davern: "Personalizing to product category knowledge: exploring the mediating effect of shopping tools on decision confidence"; Proc. Proceedings of the 37th Annual Hawaii International Conference on System Sciences, (2004), 10 pp.

[Knijnenburg et al. (2012)] B. Knijnenburg, M. Willemsen, Z. Gantner, H. Soncu and C. Newell: "Explaining the user experience of recommender systems"; User Modeling and User-Adapted Interaction, 22, 4-5 (2012), 441-504.

[Knijnenburg and Willemsen (2009)] B. P. Knijnenburg and M. C. Willemsen: "Understanding the effect of adaptive preference elicitation methods on user satisfaction of a recommender system"; Proc. Proceedings of the third ACM conference on Recommender systems, New York, New York, USA (2009), 381-384.

[Knijnenburg and Willemsen (2010)] B. P. Knijnenburg and M. C. Willemsen: "The effect of preference elicitation methods on the user experience of a recommender system"; Proc. Extended Abstracts on Human Factors in Computing Systems, Atlanta, Georgia, USA (2010), 3457-3462.

[Knijnenburg et al. (2011)] B. P. Knijnenburg, M. C. Willemsen and A. Kobsa: "A pragmatic procedure to support the user-centric evaluation of recommender systems"; Proc. Proceedings of the fifth ACM conference on Recommender systems, Chicago, Illinois, USA (2011), 321-324.

[Knotzer and Madlberger (2007)] N. Knotzer and M. Madlberger: "Consumers' Interest in Personalized Recommendations: The Role of Product-Involvement and Opinion Seeking"; Proc. 40th Annual Hawaii International Conference on System Sciences (2007), 168-168.

[Komiak and Benbasat (2006)] S. Y. X. Komiak and I. Benbasat: "The effects of personalization and familiarity on trust and adoption of recommendation agents"; MIS Quarterly, 30, 4 (2006), 941-960.

[Konstan and Riedl (2012)] J. A. Konstan and J. Riedl: "Recommender systems: from algorithms to user experience"; User Modeling and User-Adapted Interaction, 22, 1-2 (2012), 101-123.

[Kwon and Kim (2009)] O. Kwon and J. Kim: "Concept lattices for visualizing and generating user profiles for context-aware service recommendations"; Expert Systems with Applications, 36, 2 (2009), 1893-1902.

[Liang et al. (2007)] T.-P. Liang, H.-J. Lai and Y.-C. Ku: "Personalized Content Recommendation and User Satisfaction: Theoretical Synthesis and Empirical Findings"; Journal of Management Information Systems, 23, 3 (2007), 45-70.

[Mayer et al. (1995)] R. Mayer, J. Davis and D. Schoorman: "An Integrative Model of Organizational Trust"; The Academy of Management Review, 20, 3 (1995), 709-734.

[Moore and Punji (2001)] R. Moore and G. Punji: "An Investigation of Agent Assisted Consumer Information Search: Are Consumers Better Off?"; Advances in Consumer Research, 28, 1 (2001), 128.

[Olson and Widing (2002)] E. L. Olson and R. E. Widing: "Are interactive decision aids better than passive decision aids? A comparison with implications for information providers on the internet"; Journal of Interactive Marketing, 16, 2 (2002), 22-33.

[Panniello et al. (2009)] U. Panniello, M. Gorgoglione and C. Palmisano: "Comparing Pre-filtering and Post-filtering Approach in a Collaborative Contextual Recommender System: An Application to E-Commerce"; Proc. E-Commerce and Web Technologies, (2009), 348-359.

[Panniello et al. (2014)] U. Panniello, A. Tuzhilin and M. Gorgoglione: "Comparing context-aware recommender systems in terms of accuracy and diversity"; User Modeling and User-Adapted Interaction, 24, 1-2 (2014), 35-65.

[Panniello et al. (2016)] U. Panniello, M. Gorgoglione and A. Tuzhilin: "Research Note—In CARSs We Trust: How Context-Aware Recommendations Affect Customers' Trust and Other Business Performance Measures of Recommender Systems"; Information Systems Research, 27, 1 (2016), 182-196.

[Pathak and Patra (2015)] A. Pathak and B. K. Patra: "A knowledge reuse framework for improving novelty and diversity in recommendations"; Proceedings of the Second ACM IKDD Conference on Data Sciences, (2015), 11-19.

[Pathak et al. (2010)] B. Pathak, R. Garfinkel, R. Gopal, R. Venkatesan and F. Yin: "Empirical Analysis of the Impact of Recommender Systems on Sales"; J. Manage. Inf. Syst., 27, 2 (2010), 159-188.

[Pazzani and Billsus (2007)] M. Pazzani and D. Billsus: "Content-based recommendation systems."; The Adaptive Web, 4321, 2 (2007), 325-341.

[Pedersen (2000)] P. E. Pedersen: "Behavioral Effects of Using Software Agents for Product and Merchant Brokering: An Experimental Study of Consumer Decision-Making"; International Journal of Electronic Commerce, 5, 1 (2000), 125-141.

[Pereira (2001)] R. E. Pereira: "Influence of Query-Based Decision Aids on Consumer Decision Making in Electronic Commerce"; Inf. Resour. Manage. J., 14, 1 (2001), 31-48.

[Pu et al. (2011)] P. Pu, L. Chen and R. Hu: "A user-centric evaluation framework for recommender systems"; Proc. Proceedings of the fifth ACM conference on Recommender systems, Chicago, Illinois, USA (2011), 157-164.

[Pu et al. (2012)] P. Pu, L. Chen and R. Hu: "Evaluating recommender systems from the user's perspective: survey of the state of the art"; User Modeling and User-Adapted Interaction, 22, 4-5 (2012), 317-355.

[Ricci et al. (2011)] F. Ricci, L. Rokach, B. Shapira and P. B. Kantor: "Recommender Systems Handobook", (2011)

[Schafer et al. (2001)] B. Schafer, J. Konstan and J. Riedl: "E-Commerce Recommendation Applications"; Data Mining and Knowledge Discovery, 5, 1 (2001), 115-153.

[Schafer et al. (1999)] J. B. Schafer, J. Konstan and J. Riedl: "Recommender systems in e-commerce"; Proc. the 1st ACM conference on Electronic commerce, Denver, Colorado, USA (1999), 158-166.

[Schoorman et al. (2007)] F. D. Schoorman, R. C. Mayer and J. H. Davis: "An integrative model of organizational trust: Past, present, and future"; Academy of Management Review, 32, 2 (2007), 344-354.

[Sun et al. (2015)] Y. Sun, W. K. Chong, Y.-S. Han, S. Rho and K. L. Man: "Key factors affecting user experience of mobile recommendation systems"; Proc. Proceedings of the International MultiConference of Engineers and Computer Scientists, (2015),

[Swearingen and Sinha (2001)] K. Swearingen and R. Sinha: "Beyond algorithms: An HCI perspective on recommender systems"; Proc. SIGIR Workshop on Recommender Systems, New Orleans, Louisiana, USA (2001), 393-408.

[Teltzrow and Kobsa (2004)] M. Teltzrow and A. Kobsa: "Impacts of User Privacy Preferences on Personalized Systems"; Designing Personalized User Experiences in eCommerce, 5, (2004), 315-332.

[Venkatesh and Davis (2000)] V. Venkatesh and F. D. Davis: "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies"; Manage. Sci., 46, 2 (2000), 186-204.

[Venkatesh et al. (2003)] V. Venkatesh, M. G. Morris, G. B. Davis and F. D. Davis: "User Acceptance of Information Technology: Toward a Unified View"; MIS Quarterly, 27, 3 (2003), 425-478.

[Vijayasarathy and Jones (2001)] L. R. Vijayasarathy and J. M. Jones: "Do Internet Shopping Aids Make a Difference? An Empirical Investigation"; Electronic Markets, 11, 1 (2001), 75-83.

[Wang et al. (2014)] Y.-Y. Wang, A. Luse, A. Townsend and B. Mennecke: "Understanding the moderating roles of types of recommender systems and products on customer behavioral intention to use recommender systems"; Information Systems and e-Business Management, (2014), 1-31.

[Xiao and Benbasat (2007)] B. Xiao and I. Benbasat: "E-commerce product recommendation agents: use, characteristics, and impact"; MIS Quarterly, 31, 1 (2007), 137-209.

[Ziegler et al. (2005)] C.-N. Ziegler, S. M. McNee, J. A. Konstan and G. Lausen: "Improving recommendation lists through topic diversification"; 14th international conference on World Wide Web, (2005), 22-32.