Electronic Commerce Research and Applications 19 (2016) 70–84

Contents lists available at ScienceDirect

Electronic Commerce Research and Applications

journal homepage: www.elsevier.com/locate/ecra

Recommendation quality, transparency, and website quality for trust-building in recommendation agents

Mehrbakhsh Nilashi [a,c,*], Dietmar Jannach [b], Othman bin Ibrahim [a], Mohammad Dalvi Esfahani [a], Hossein Ahmadi [a]

[a] Faculty of Computing, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
[b] TU Dortmund, 44221 Dortmund, Germany
[c] Department of Computer Engineering, Lahijan Branch, Islamic Azad University, Lahijan, Iran

ARTICLE INFO

Article history:
Received 30 October 2015
Received in revised form 3 September 2016
Accepted 3 September 2016
Available online 8 September 2016

Keywords:
Empirical study
Recommendation agents
Survey research
Transparency
Trust
Website quality

ABSTRACT

Trust is a main success factor for automated recommendation agents on e-commerce sites. Various aspects can contribute to the development of trust toward such an agent, including perceptions about the usefulness of the recommendations, the transparency of the recommendation process, and the general quality of the website. These factors have been analyzed in isolation in the literature though. We propose and evaluate a new trust model that integrates these factors, and allows us to assess their relative importance for trust-building. We conducted empirical studies in the context of two popular e-commerce websites. The findings suggest that transparency is equally important to consumers for building trust as recommendation quality, and that general we site quality contributes to the development of trust. The findings indicate that focusing on recommendation quality may be insufficient and higher levels of adoption of the recommendations can be achieved when several trust-building factors are considered.

1. Introduction

Automated recommendations have become a ubiquitous part of our online user experience and many modern e-commerce providers employ recommendation agents on their websites. The task of such agents is to point consumers to additional items of interest in the catalog, either based on their current activity or based on their long term preference profile. Several works in the literature show that recommendation agents have a measurable impact on the consumers' purchase behavior and can also improve their decision-making processes (Aksoy et al., 2006; Gomez-Uribe and Hunt, 2015; Häubl and Murray, 2006; Häubl and Trifts, 2000; Xiao and Benbasat, 2007).

Several studies in real-world environments provide evidence that recommendation agents can drive the short-term behavior of their customers (Dias et al., 2008; Garcin et al., 2014; Jannach and Hegelich, 2009; Zanker et al., 2006). Providers are typically interested in the long-term success of their websites and the embedded recommendation agents though. Whether consumers adopt the recommendations in the long run depends on their trust in the recommendation agent (Grabner-Kräuter and Kaluscha, 2003), which can be developed by repeated positive experiences.

A key prerequisite for such positive experiences and, consequently, trust is that the recommendation agent is continuously able to generate useful recommendations that help the consumers make better decisions or find better suited items in the catalog. Improving the quality of the recommendations in different dimensions is therefore a major focus in the academic literature on recommender systems (Jannach et al., 2012). In addition to the capability of an agent to select presumably relevant items for a consumer, the perceived transparency of the recommendation process is often considered a key factor for the establishment of trust toward the recommendation agent (Gedikli et al., 2014; Sinha and Swearingen, 2002; Tintarev and Masthoff, 2007b). Transparency means that the recommendation agent is capable of conveying to the user – for example with the help of system-generated explanations – why certain items were recommended.

Looking beyond trust-related aspects of recommendation agents, the topic of trust in e-commerce has been extensively discussed in the research literature, and trust is often considered as a key enabler for the success of e-commerce in general (Adamopoulou and Symeonidis, 2014; Gefen, 2000; Grabner-Kräuter and Kaluscha, 2003; Jarvenpaa et al., 1999; Moody et al., 2014). In the e-commerce settings considered in this paper, the online shop is often the only way in which a company communicates with its customers (Chang and Chen, 2008), and one common assumption in this stream of research is that Internet users at least

* Corresponding author.
E-mail address: nilashidotnet@hotmail.com (M. Nilashi).

http://dx.doi.org/10.1016/j.elerap.2016.09.003
1567-4223/

partially rely on website attributes like information quality, interaction quality, or performance to judge the trustworthiness of the site. Since recommendation agents are typically part of the web shop, we can conjecture that such website quality attributes may also affect to what extent consumers trust the embedded agent. Overall, a number of aspects can contribute to the development of trust toward a recommendation agent (Wang and Benbasat, 2008). A number of previous studies focused on the understanding of individual components of trust-establishment. Pu and Chen (2006, 2007) and Zanker (2012), for example, focus mainly on the role of explanations as a mechanism for trust development. Komiak and Benbasat (2006) on the other hand found that personalization positively affects the trust toward the agent.

Continuing these lines of research on trust-promoting factors, the goal of our work is to investigate and quantify the relative importance of the different factors. Our general research question is: What is the relative importance of general website quality, recommendation quality, and transparency as trust-building factors for recommendation agents in e-commerce?

To answer this question, we developed a layered trust model that considers the following: recommendation quality, transparency of the recommendation process, and general website quality. The trust-promoting factors recommendation quality and website quality are furthermore decomposed into components in the layered model, which allows us to obtain a better understanding of the role of their constituent factors. Since from a provider's perspective trust is not the ultimate variable of interest, our trust model, unlike for example the one proposed in (Wang and Benbasat, 2008), includes the analysis of the impact of trust on the user's behavioral intentions, which is in our case their intention to adopt the recommendations made by an agent.

Overall, the novelty of our model lies in its comprehensive nature, layered design that allows us to assess the relative importance of the factors, and the consideration of website quality as a trust-building factor, which has not been investigated in combination with the other aspects. In contrast to previous works (Baier and Stüber, 2010; Pu and Chen, 2006; Wang and Benbasat, 2008), which base their empirical evaluations on websites that were only created for the purpose of the studies, we conducted our analyses with two real e-commerce websites.

As a practical outcome, our findings should help the providers of recommendation services make better-informed decisions about which aspects they should put particular focus on when trying to improve the adoption of the recommendation agent.

2. Theoretical background

Our proposed trust model combines existing insights related to general trust-promoting factors in e-commerce as well as specific trust-building factors for recommendation agents.

2.1. Trust in e-commerce systems and the WebQual model

We first briefly review the literature on the general role of trust in e-commerce settings and elaborate on the WebQual model, which serves as a basis for the development of the website quality aspect of our integrated trust model.

2.1.1. The role of trust in e-commerce

The question of which factors help to promote the consumers' trust toward online stores or e-commerce websites has been extensively covered in the fields of information systems (IS), marketing, and e-commerce. One frequently cited definition of trust in various contexts is the "willingness to be vulnerable," proposed by Mayer et al. (1995). Chopra and Wallace (2003) define trust in the e-commerce environment as the willingness to rely on a specific other, based on the confidence that one's trust will lead to positive outcomes. For Lim et al. (2006), consumer trust in online shopping is defined as the willingness of a consumer to be exposed to the possibility of loss during an Internet shopping transaction. This trust is based on the expectation that the merchant will engage in generally-acceptable practices, and will be able to deliver the promised products or services.[1]

As the definitions indicate, consumer trust is a precursor of behavioral intentions (Gefen et al., 2003; Pu et al., 2011), referred to as "trusting intentions" (McKnight et al., 1998). Different empirical studies have shown that trust in an e-commerce website can help to increase the customers' intention to purchase products from it (Corbitt et al., 2003).

A number of trust-promoting factors and their impact on consumer behavior intentions have been explored, including for example, the perceived level of privacy and security (Pavlou and Chellappa, 2001), organizational reputation, relative advantage and perceived risk (de Ruyter et al., 2001). The relation between the perceived size and perceived reputation of an Internet store and its impact on trust was discussed in Jarvenpaa et al. (1999). Karimov et al. (2011) focused on establishing trust. Their review provides ample support that website design is important for developing trust in unfamiliar online vendors and emphasizes the potential value of recommendation agents.

The effects of visitors' experiences with a user interface on trust formation have also been investigated (Koufaris and Hampton-Sosa (2002). In their work, a trust model was established and evaluated which showed that usefulness and perceived ease of use of the website are positively associated with consumer trust in the online company and the customers' intentions to purchase and intentions to return. The role of a website's appeal and usability on trust was further explored (Hampton-Sosa and Koufaris, 2005). And finally, a model was proposed that included a familiarity and trust aspect of e-commerce adoption (Gefen et al., 2003). The related analysis showed that purchase intentions of repeated buyers were influenced by the customer's trust in the e-vendor and the perceived usefulness of the website.

2.1.2. The WebQual model

Since the importance of website quality is a general trust-promoting factor, a methodology is needed to measure it. We rely on the WebQual model (Barnes and Vidgen, 2002, 2003), which was s designed to evaluate and operationalize this measurement. WebQual numerically captures the quality of a website but is based on subjective quality impressions. So it is suitable for our research, which is also based on subjective quality perceptions.

With WebQual, the quality of a website can be expressed in terms of numerical quality indicators in different dimensions with the help of a questionnaire. The relevant dimensions are related: (1) information quality as a core area of IS research; (2) interaction and service quality from marketing, e-commerce, and IS service quality research; and (3) usability aspects from the field of Human Computer Interaction. The three major dimensions of e-commerce website quality according to the WebQual 4.0 model are usability, information quality, and service interaction quality.

For each of the dimensions, Barnes and Vidgen (2002) propose a number of questionnaire items (e.g., "I find the site easy to navigate" for the usability dimension). In our work, we adopt the quality factors and the related questionnaire items of the established WebQual model to assess the perceived quality of the website. In our case, we are not interested in the absolute

[1] A detailed review of the characteristics of digital trust, the factors that influence trust, and approaches to trust modeling and management can be found in Yan and Holtmanns (2007).

values of the individual quality indicators but rather in their relationships with the other variables of our integrated trust model.

Alternative frameworks for assessing the quality of online services and Websites are described by several authors (Hasan and Abuelrub, 2011; Jiang et al., 2002; Olsina and Rossi, 2002; Wang et al., 1999). Some of them are however more complex than WebQual or have to be tailored to our specific problem setting first. Note that website quality may not only influence the consumers' trust as proposed in our research. Website quality aspects have also been shown to be relevant factors that can influence the general online shopping value and customer repurchase intentions (Huang and Benyoucef, 2013; Kim et al., 2012).

2.2. Trust in recommendation agents

Next, we review existing works on trust-promoting factors that are specific to recommendation agents. We discuss those factors that are part of our proposed trust model, especially recommendation quality, as well as transparency and explanations.

2.2.1. Trust-building factors of recommendation agents

A variety of aspects can contribute to the adoption and success of e-commerce recommendation agents. Xiao and Benbasat (2007) developed a comprehensive conceptual framework and made 28 propositions based on existing research on how different factors impact the users' decision making processes and their evaluation of a recommendation agent. In the context of trust, which they define as the "user beliefs in the [agent's] competence, benevolence, and integrity," the authors identified that, among others, the type of the recommendation agent – pure, content-based, or hybrid – the reputation of the provider, and the ease of use of the system can increase the consumers' trust in the system. In another study, Wang and Benbasat (2008) investigated the main reasons why users trust or do not trust a decision support technology during the early stage of using it and highlight that trust toward a recommendation agent is mainly triggered by the experiential use of the system.

Besides these factors, user experience, user control, and the usability of the system were mentioned as trust-building components in other research on recommender systems (Konstan and Riedl, 2012; McNee et al., 2006; Polatidis and Georgiadis, 2014; Sinha and Swearingen, 2001; Swearingen and Sinha, 2002). More recently, user experience in mobile recommender systems (Sun et al., 2015a,b) has been examined, for additional factors related to pull versus push communication and privacy issues.

Our proposed research model covers such aspects of usability and user experience in general through the constructs that are related to website quality. Also different other aspects mentioned in the literature – such as the reputation of the provider, the perceived level of personalization, or the provided information – are part of our model through usability, service interaction, and information quality which we adopted from the WebQual model.

2.2.2. Recommendation quality: accuracy, novelty, and diversity

Whether online users will trust the advice of a recommendation agent in the long run and continue to adopt the recommendations depends on the usefulness of the recommendations, especially their quality. What makes a recommendation useful, may again depend on different factors.

The main research focus in computer science is on improving a recommendation agent's ability to predict to what extent an item will match the user's preferences and interests. This is accuracy (Herlocker et al., 2004), and our model includes a related construct, recommendation accuracy. Being able to accurately predict the user's preference does not directly inform us to what extent such recommendation approaches are suited to increasing consumer trust in the recommendations. Indirectly, being able to accurately predict the preferences allows us to achieve more precise forms of personalization and higher user satisfaction. These may result due to better decision support (Aksoy et al., 2006; Liang et al., 2007; Ku and Tai, 2013; Komiak and Benbasat, 2006). This, in turn, helps to establish long-term trust toward the recommendation agent. On the other hand, if the users feel that the recommendations are unsuited or biased, they can begin to distrust the system, which in the worst case has a negative impact on the website's performance (Chau et al., 2013).

High levels of accuracy allow us to distinguish between items that are relevant for the user and those that are not. Depending on the domain, focusing only on accuracy may not be enough to generate useful recommendations (McNee et al., 2006). Consider, for example, a recommendation agent for an online music service, which recommends tracks to play to its users. If the agent mostly recommends tracks of an artist that the user has listened to in the past, the novelty of the recommendations for the user may be limited. It will not help the user much in discovering new artists or musical styles, a typical expectation of users of such services in some domains.[2]

Vargas and Castells (2011) propose different computational metrics to assess the level of novelty of recommendations. Their metrics, for example, consider the general popularity of an item in the community or the similarity of a recommendable item to the items that the user had positive experiences with in the past. In addition to these factors, Ge et al. (2010) consider serendipity as a form of novelty that has the potential to contribute to the usefulness of the recommendations. Our trust model thus includes a construct, recommendation novelty, which captures these aspects of recommendation quality. Note that higher levels of novelty do not necessarily lead to better perceived recommendation quality (Cremonesi et al., 2011; Ekstrand et al., 2014; Said et al., 2013), when the recommended items are too different from the user's general taste and the user's familiarity with the recommended items is low.

In addition to novelty, diversity has received increased interest as a quality factor in recent years. The underlying assumption is that, if the recommended items are too similar too each other, the value for the user can be limited. If all recommended items are quite similar, they may not be well-suited to help users to understand the space of options or may be perceived as being too monotonous or even biased. In fact, limited recommendation diversity and catalog coverage can be problematic also from the perspective of the provider of the recommendation service (Bodoff and Ho, 2015; Fleder and Hosanagar, 2009). Recommendation diversity is typically measured in the literature in terms of the similarity of the recommended items to each other and different studies suggest that increasing the level of diversity has a positive effect on the users' quality perception of the recommendations (Ekstrand et al., 2014; Said et al., 2013; Ziegler et al., 2005). We thus include the corresponding construct recommendation diversity in our trust model.

Transparency, an agent's capability of explaining to the user why certain items were recommended, is viewed as a central trust-building factor (Pu and Chen, 2007; Swearingen and Sinha, 2002; Wang and Benbasat, 2005; Xiao and Benbasat, 2007). We include transparency in our research model. Similar to the other factors, we consider how users perceive the transparency of the recommendation process, and how users think that the agent determines the recommendations. Historically, explanation facilities as trust-building mechanisms have been in the focus of

[2] Depending on the domain, recommending items that are already known to a user can also be valuable. See Lerche et al. (2016) on the value of reminders in e-commerce recommendations.

research in different fields already, for example, in traditional knowledge-based expert systems and general IS (Johnson and Johnson, 1993). In the context of recommendation agents, early studies by Swearingen and Sinha (2002) showed the positive role of transparency for establishing trust. Another example of trust-related research is Zimmerman and Kurapati (2002), who proposed a method for increasing trust in a TV show recommender, which was based on displaying a reflective history in the user interface of the system.

Questions of how to present the explanations to users and which information to display have been explored too (Gedikli et al., 2014; Tintarev and Masthoff, 2007a; Herlocker et al., 2000). The focus was not on trust though. Instead, the studies focus on individual effects of explanations such as perceived transparency, effectiveness, efficiency, and persuasiveness. These serve as a basis for the development of trust toward the agent. The studies also emphasize the importance of the interface in transparency and trust.

Trust-inspiring explanation interfaces have also been explored (Pu and Chen, 2007). The authors develop a research model to analyze the effects of two different user interfaces on user trust toward the agent. Their work is similar to our research in that we aim to identify and quantify trust-building factors. We do not compare artificial user interfaces but instead use real online shops. Further, our research approach includes trust as a factor in the model.

Recently, Wang et al. (2016) investigated the differential effects of using avatar interfaces and explanation facilities as a means of increasing the persuasive power of recommendation agents. Observations from an experimental evaluation confirm that explanations are suited to increasing cognitive, rationality-based trust. Interface avatars, on the other hand, have the potential to create affect-based, emotional trust, provided that the implementation of the avatar is considered as being highly professional by the participants. In our research model, we focus on cognitive trust; the real-world websites used in our experiments do not use interface avatars.

3. Research model and hypotheses

Our review of the literature shows that various factors contribute to the user's trust toward a recommendation agent. Our resulting research model is based on the three main factors: general website quality, recommendation quality, and transparency. The key research question is about the relative importance of the factors. Fig. 1 shows the hypothesized effects among the variables in our layered model. We discuss our hypotheses in detail. We measure the different qualities based on the responses of participants in two user surveys. All measurements relate to perceived qualities. We do not measure recommendation quality with objective measures such as precision and recall though.

3.1. Website quality and trust: Hypotheses 1–4

The WebQual model assumes that the quality of a website is influenced by its usability, information content, and aspects of service interaction. These directly lead us to the first three research hypotheses related to website quality. The literature confirms in various forms that aspects of website quality have an impact on the consumers' trust toward the website. Based on these findings, we also offer Hypothesis 4, that website quality has a positive trust-inspiring effect toward the recommendation agent.

To summarize:

- **Hypothesis 1 (Usability → Website Quality)**. *The usability of a website that hosts a recommendendation agent has a positive effect on perceived website quality.*
- **Hypothesis 2 (Information Quality → Website Quality)**. *Information quality has a positive effect on perceived website quality.*
- **Hypothesis 3 (Service Interaction Quality → Website Quality)**. *Service interaction quality has a positive effect on perceived website quality.*
- **Hypothesis 4 (Website Quality → Recommendation Trust)**. *Website quality has a positive effect on consumer trust in the recommendation agent.*

3.2. Recommendation quality and trust: Hypotheses 5–8

Several factors, accuracy, novelty, and diversity, may impact the user's perception of a recommendation agent's quality. Not all effects are positive, and Ekstrand et al. (2014) and Jannach et al. (2015) suggest that a high level of novelty can be detrimental to consumer perceptions of quality. These observations lead us to

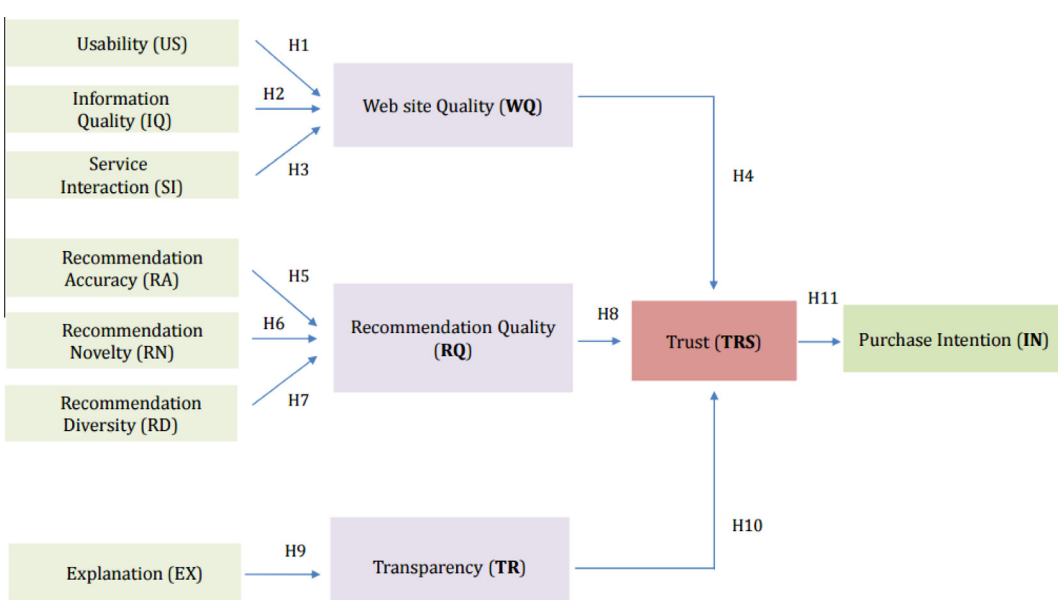

Fig. 1. Trust model for e-commerce recommendation agents.

the Hypotheses 5–8. We relate these quality factors to recommendation quality. In Hypothesis 8, we assert that recommendation quality has an influence on the consumer trust because of a higher perceived level of personalization. Again, to summarize:

- **Hypothesis 5 (Recommendation Accuracy → Recommendation Quality)**. *The accuracy of the recommendations has a positive effect on perceived recommendation quality.*
- **Hypothesis 6 (Recommendation Novelty → Recommendation Quality)**. *Recommendation quality has no positive effect on perceived recommendation agent quality.*
- **Hypothesis 7 (Recommendation Diversity → Recommendation Quality)**. *Recommendation diversity has a positive effect on perceived recommendation agent quality.*
- **Hypothesis 8 (Recommendation Quality → Trust in the Recommendation Agent)**. *Perceived recommendation quality has a positive effect on consumer trust in the recommendation agent.*

3.3. Transparency, trust, and intention to purchase: Hypotheses 9–11

The final three hypotheses concern the relationships between the explanations made by the recommendation agent, the perceived level of transparency, and the possible effects on trust and the user's intention to make a purchase (Pu et al., 2011). The hypotheses are summarized as follows:

- **Hypothesis 9 (Explanations → Recommendation Process Transparency)**. *Explanations have a positive effect on the perceived transparency of the recommendation process.*
- **Hypothesis 10 (Recommendation Process Transparency → Recommender System Trust)**. *Transparency has a positive effect on the consumers' trust in the recommendations made by the agent.*
- **Hypothesis 11 (Trust → Purchase Intention)**. *Increased levels of trust positively impact a consumers' intention to purchase an item.*

4. Survey design

We conducted a laboratory study with 150 participants to test our research hypotheses. The survey design and our observations are described next.

4.1. Participants

We recruited 150 master and PhD students Computer Science and Information Systems at the Universiti Teknologi Malaysia (UTM) to participate in the survey. Most of them were between 23 and 35 years old and experienced users of online shopping platforms. The detailed statistics of the students, who were selected based on convenience sampling, are given in Table 1.

Table 1
Demographic profile of the participants.

	Items	Frequency	Percent
Age	22	14	9.33
	23–26	44	29.33
	27–35	54	36.00
	36–40	23	15.33
	>40	15	10.00
Gender	Male	87	58.00
	Female	63	42.00
Education	Master stud.	49	32.67
	Ph.D. stud.	101	67.33
Experience with Amazon.com	<1 Years	3	2.00
	1–3 Years	9	6.00
	>3 Years	138	92.00
Experience with Lazada.com	<1 Years	12	8.00
	1–3 Years	17	11.33
	>3 Years	121	80.66

4.2. Data collection process

The task of the participants was to interact with two different e-commerce websites (Amazon.com and Lazada.com). On Amazon.com, they were asked to inspect and select 10 different shop items in the book section of the online store. On Lazada.com, the target domain was digital cameras. Choosing two different item categories should allow us to validate that the observed effects generalize to different domains.

When visiting the Amazon.com website, participants could look up books of their choice but were advised to focus on books they already know about (e.g., technical books that they have read during their studies). Comparable instructions were made for the camera domain on Lazada.com. The participants also were encouraged at the beginning of the experiment to look at the recommendations provided on the site when they were browsing the store and to inspect the details of the books or cameras.

As shown in Table 1, the majority of the participants was already well-acquainted with both systems before the experiment. For example, more than 90% had been using Amazon.com for more than three years. Even for the much younger Lazada platform, which has a similar business model as Amazon, more than 80% had been using the e-commerce website beginning with or shortly after its introduction. The experiment was conducted in one of the laboratories of our university.

All of the computer that were used had the same hardware equipment in terms of screens, memory, and processing power. On average, the participants spent about 90 min interacting with both websites and inspecting the item details. Instead of actually buying books or cameras, the participants could place relevant items into the shopping basket. At the end of the browsing session, the participants were asked to fill out a questionnaire.

4.3. Questionnaire

The questionnaire comprised 48 items related to the 12 constructs of our research model. The question items are listed in Table 2 and are organized in the following groups. (See the Appendix Table A1 for the references for the questionnaire items.)

- Usability (US1–US8), Information Quality (IQ1–IQ7), Service Interaction (SI1–SI7): The question items are based on the WebQual model.
- Recommendation Accuracy (RA1–RA3), Recommendation Novelty (RN1–RN3), Recommendation Diversity (RD1–RD2): Questions relate to the perceived quality of the recommendations. The question items are similar to those suggested and used previously (Ekstrand et al., 2014; Knijnenburg et al., 2012; Pu et al., 2011).
- Explanations (EX1–EX3): The question items are inspired by the questionnaire of Gedikli et al. (2014) and the considerations of Herlocker et al. (2000) about explanations and user involvement.
- Website Quality (WQ1–WQ3), Recommendation Quality (RQ1–RQ3), Transparency (TR1–TR3): The questionnaire items assess the users' perception of the different quality factors.
- Trust (TRS1–TRS3): The three questionnaire items assess the users' trust in the system.
- Purchase Intention (IN1–IN3): The final questionnaire items were designed to gauge the users' intention to purchase an item.

Table 2
Questionnaire items.

Construct	Item
Usability (US)	1. I find the site easy to learn to operate
	2. My interaction with the site is clear and understandable
	3. I find the site easy to navigate
	4. I find the site easy to use
	5. The site has an attractive appearance
	6. The design is appropriate for this type of website
	7. The site conveys a sense of competence
	8. The site creates a positive experience for me
Information Quality (IQ)	1. The website provides accurate information
	2. The website provides believable information
	3. The website provides timely information
	4. The website provides relevant information
	5. The website provides easy to understand information
	6. The website provides information at the right level of detail
	7. The website provides the information in an appropriate format
Service Interaction (SI)	1. The site has a good reputation
	2. It feels safe to complete transactions
	3. My personal information feels secure
	4. The site creates a sense of personalization
	5. The website conveys a sense of community
	6. It is easy to communicate with the organization
	7. I feel confident that goods/services will be delivered as promised
Web site Quality (WQ)	1. My overall evaluation of the features of this website is very high
	2. The quality of this Web site meets my expectations
	3. The Web site offered unique features to me that are different from other retail Web sites
Recommendation Accuracy (RA)	1. The items recommended to me matched my interests
	2. I find many items appealing that the system recommended me
	3. The recommendations I received better fits my interests than what I may receive from a friend
Recommendation Novelty (RN)	1. This recommendation agent recommended items to me that I did not expect
	2. This recommendation agent helped me discover new products
	3. I could find familiar items through the recommender
Recommendation Diversity (RD)	1. The items recommended to me are not similar to each other
	2. The items recommended to me are of various kinds
Recommendation Quality (RQ)	1. The quality of the recommendations is the same as I wanted
	2. My overall evaluation of the recommendations of this recommendation agent is superior
	3. Overall, the quality of the recommendations of this recommendation agent valuable
Explanation (EX)	1. The recommender explains why the products are recommended to me
	2. When interacting with the recommendation agent, I felt involved in its recommendation process
	3. This recommendation agent educated me about the process used for generating a recommendation, so that I could better understand the strengths and limitations of the system
Transparency (TR)	1. I understood why the items were recommended to me
	2. The explanation facilities of the recommendation agent help me make better decisions
	3. The explanation facilities helped increase my acceptance of the recommendations made by the system
Trust (TRS)	1. I am convinced that the recommended items are suitable for me
	2. I am confident I will like the items recommended to me
	3. The recommender can be trusted
Purchase Intention (IN)	1. I would buy the items recommended, given the opportunity
	2. I have a high intention to buy items recommended by the recommender
	3. I intend to continue using this recommender for purchasing items in the future

The participants had to express their opinions using five-point Likert scales (Strongly Agree = 5, Agree = 4, Neither Agree nor Disagree = 3, Disagree = 2, and Strongly Disagree = 1).

4.4. Differences between Amazon.com and Lazada.com

The business models of Amazon.com and Lazada.com, which are popular in Southeast Asia, are comparable: they both offer a broad range of items on their e-commerce websites. Amazon.com pioneering the use of recommendations on a large scale (Linden et al., 2003), and provides recommendations in a large number of item categories, typically under the labels "Customers who bought . . . also bought" or "Frequently bought together." On Lazada.com, comparable recommendations are presented in many categories under the labels "Other options", "Related products", "Frequently bought together", or "Related searches." In both cases, the presented items are neither explicitly labeled as personal recommendations, nor do the labels suggest that the recommendations are personalized. In our view, their functionality is comparable.

Regarding the aspects of explanations and transparency, note that the website of Amazon.com – at the time of the study – had different features suited to explain the rationale of the recommendations to the user. On Amazon.com the different labels of the various recommendation lists carry some explanatory information. Furthermore, inside certain recommendation lists, each item has a link attached labeled "Why recommended." Clicking on the link opens a pop-up window in which an explanation is presented that is usually based on other purchases or item views. Within the pop-up window, users can rate the recommended item and the items used in the explanations. Further, they can give feedback to the system to not use certain items as a basis for recommendations anymore.

Such functionality is not available on Lazada.com; the main explanatory information is conveyed through the choice of the labels of the recommendation lists (Gedikli et al., 2014). Both e-commerce websites prominently display information about customer ratings and reviews on each item page to make the overall rating of the presented items transparent for the users though.

We focus on the concept of perceived transparency, the subjective feeling of the users that they understand the recommendations. Past research has shown that it is not necessarily required that the explanations truly reflect the inner workings of a recommendation system to be able to have a persuasive effect on users (Herlocker et al., 2000).

5. Empirical results

We applied partial least squares (PLS) and structural equation modeling (SEM) (Hair et al., 1998) to analyze the hypothesized relationships in our research model. SEM is a quantitative statistical method to assess the relationships between multiple independent variables and dependent variables of a model. It combines the advantages of path analysis, factor analysis, and multiple regression analysis and helps us examine the relationship between the variables in terms of the explained variance (Jöreskog and Sörbom, 1993). We used SmartPLS (www.smartpls.com) for our analysis.

Before we summarize the results for our research model, we will first report the outcomes of different tests that we applied to check the appropriateness of the model.

5.1. Assessment of the measurement model

We assess the reliability and validity of our measurement model in four different dimensions: indicator reliability, internal consistency, convergent validity, and discriminant validity (Hair et al., 2013; Henseler et al., 2009).

5.1.1. Indicator reliability

Indicator reliability refers to how well each item is related to its respective construct (Hair et al., 2013). Indicators should have outer loadings above the threshold value of 0.7. In case the values are below the threshold value of 0.7, the indicators should be considered for removal (Hair et al., 2011). In our research, we adopted 0.7 as an acceptable level for item loading. The results for indicator reliability are shown in Tables A.1 (Amazon case) and A.2 (Lazada) in the Appendix.

All outer loadings of the reflective constructs are well above the threshold value of 0.7. All 48 items were above the minimum acceptable level for the outer loadings for both e-commerce websites that we tested. In the case of Amazon, the indicator US2 (0.729) – about the clarity of the interaction with the site – had the smallest indicator reliability and the indicator RD2 (0.931) – on recommendation diversity – had the highest one. In the Lazada case, the lowest value was observed for RA1 (accuracy) and the highest value again for RD2.

5.1.2. Internal consistency, convergent validity and discriminant validity

The reliability of each specified research construct was checked before performing the validity analysis. In Appendix Tables A.3 (Amazon) and Table A.4 (Lazada), we report the values for composite reliability, which range from 0.834 to 0.963, and Cronbach's alpha, with values between 0.702 and 0.956. All values are higher than the suggested minimum threshold of 0.7, thus representing an adequate level of internal consistency (Sekaran and Bougie, 2010; Nunnally, 1978).

Convergent validity refers to whether a latent variable is able to explain its indicators on average; it signifies the amount of variance shared between a construct and its indicators. Convergent validity builds on the average variance extracted (AVE) value. Following Hair et al. (2013), we consider 0.5 as the acceptable minimum AVE value. As shown in Appendix Table A.5, the AVE values for all constructs exceeded the recommended value of 0.5, which means that convergent validity is established for both e-commerce websites. The highest values for AVE were found for recommendation diversity for both, and the lowest acceptable values were found for usability for Amazon and for recommendation accuracy for Lazada.

To assess the discriminant validity of the model, we used the cross loadings matrix and the Fornell–Larcker criterion (Fornell and Larcker, 1981). In the cross-loadings matrix, the item loading of each indicator should be greater than all of its loadings on other constructs. According to the Fornell–Larcker criterion, the square root of the AVE of each construct must be higher than its highest correlation with any other construct. The logic of this method is that a construct shares more variance with its associated indicators than with any other construct (Hair et al., 2013).

Appendix Tables A.1 (Amazon) and A.2 (Lazada) also show the cross-loadings matrices. The results reveal that the loading of each measurement item on its assigned construct is larger than its loading on any other construct in the model. Thus, the items for each construct were loaded on a single factor without large cross-loadings on additional factors. The results confirm that the measurement model has strong discriminant validity at the level of the items. The results for the discriminant validity analysis based on the Fornell–Larcker criterion are shown in Appendix Tables A.6 (Amazon) and A.7 (Lazada). When applying this criterion, the square root of the AVE of each construct is compared with its bivariate correlations with all other constructs. Thus, we inspected the square root of the AVEs to ensure that their values were above the correlations between the various constructs.

The results in Tables A.6 (Amazon) and A.7 (Lazada) show that in all cases the square root of the AVE of each construct is larger than the inter-construct correlations (Fornell and Bookstein, 1982). In addition, an analysis of the correlations among all latent variables shows that all constructs are distinct from each other because their correlations are below the threshold of 0.9.

5.1.3. Assessment of the formative construct

Information quality is the only formative construct of our model, and all the others are reflective constructs. To ensure the validity and reliability of the construct, we check possible collinearity issues and the construct's significance and relevance (Hair et al., 2011; Wilcox et al., 2008).

To check for collinearity, the variance inflation factor (VIF) was assessed. High correlation between two formative indicators can indicate a possibly problematic collinearity issue. Following Hair et al. (2013), a VIF of 5 and higher indicates the existence of collinearity. The results of the analysis are given in Appendix Table A.8. All VIF are below the threshold of 5 for both e-commerce websites that we tested.

To assess the significance of the path coefficients, we use the bootstrapping algorithm. The outer weights obtained by the bootstrapping procedure should be different from zero and surpass a minimum threshold 1.96 for the *t*-value. In the Amazon case, all outer weights are different from zero, but only IQ2 meets the minimum threshold for the *t*-value. In such cases, as suggested by Hair et al. (2013), the outer loadings should be checked for these particular indicators to see if they pass a minimum threshold of 0.50. If so as it is in our analysis for the Amazon site, they should be retained in the analysis. For the Lazada site, all measurement items except IQ7 met the minimum threshold of 0.5 for the outer loading. So we removed IQ7 from the subsequent analysis.

The detailed statistical results are presented in Appendix Tables A.9 and A.10.

5.1.4. Summary of the analysis of the measurement model

Our assessment of the measurement model shows that there is satisfactory empirical support with respect to reliability as well as

convergent and discriminant validity. The assessment of the formative construct information quality revealed that measurement item IQ7 did not reach the required significance level and was removed from the further analysis in the case of the Lazada site. For the Amazon case, all of the 48 measurement items were retained. With an adequate and sufficient measurement model, we next present the assessment of our structural model using the PLS and SEM analysis.

5.2. Assessment of the structural model and research hypotheses

The structural or inner model represents the relationship between the variables and constructs. We use two criteria to assess the structural model: the size and significance of the path coefficients and the coefficient of determination R^2 (Hair et al., 2013; Kijsanayotin et al., 2009; Urbach and Ahlemann, 2010).

5.2.1. Path coefficients

Path coefficients indicate the strength of the relationship between the latent variables in the model and range between −1 and +1. The *t*-values obtained by bootstrapping evaluate whether a coefficient is significant. The obtained path coefficients and *t*-values are listed in Tables 3 (Amazon) and Table 4 (Lazada).

The main result is that for both e-commerce websites all three main factors proposed in our model (website quality, recommendation quality, and transparency) significantly affect the consumers' trust in the system ($p < 0.01$). Trust, in turn, has a significant impact on the users' purchase intention ($p < 0.01$). And most of the other hypothesized effects are also significant ($p < 0.01$). In two cases, the effects were only marginally significant for one of the e-commerce websites ($p < 0.10$) even though they were clearly significant ($p < 0.01$) in the other case. Both relate to the Usability → Website Quality Hypothesis (H1) and the Service Interaction Quality → Website Quality Hypothesis (H3) on the relationships between usability, service interaction, and website quality. That these relationships were only marginally significant for one of the websites is surprising as the hypotheses were based on the WebQual model.

Regarding the different quality factors for recommendations, novelty – in contrast to the Recommendation Novelty → Recommendation Quality Hypothesis (H6) – had the strongest positive effect on the perceived quality of the recommendations for both applications. To some extent this is surprising, since in other domains like movie recommendation novelty had a limited effect on the user's perception of the system (Ekstrand et al., 2014; Jannach et al., 2015). The importance of the quality factors seems to partially depend on the domain, as the importance of diversity was inconsistent across the two websites.

5.2.2. Coefficients of determination (R^2s)

The second key criterion for evaluating the structural model is the coefficient of determination R^2 for the dependent variables. It measures the proportion of the variance of a dependent variable that is explained by independent variables (Hair et al., 1998). It shows the model's ability to explain and predict the dependent latent variables (Ringle et al., 2012). R^2 values of 0.75, 0.50, or 0.25 for dependent variables are viewed as substantial, moderate, or weak (Hair et al., 2013).

Tables 5 (Amazon) and Table 6 (Lazada) present the R^2s of the dependent variables in the research model. Effects could be observed for all dependent variables. Most measured effects are above the moderate threshold. The only exception for both e-commerce websites concerns the relationship for trust and purchase intention, where the values are below 0.5. So the effects are weaker than the others, but still present.

5.2.3. Summary of model fit results

The overall results for our proposed model are summarized in Figs. 2a (Amazon) and 2b (Lazada).

The results for the two e-commerce websites show that:

- Trust in both cases has a strong impact on the user's intention to make a purchase with β values being around 0.6.
- Transparency, recommendation quality, and website quality all have significant effects on trust too. The relative importance of the factors is consistent for both websites as well.
- Explanations lead to comparable effects on transparency with values between 0.7 and 0.8 in both.

Table 3
Evaluation of research hypotheses (Amazon).

Hyp.	Link	β	t-value	p-value
H1	US → WQ	0.125	1.847	0.065
H2	IQ → WQ	0.319	3.982	0.000
H3	SI → WQ	0.393	5.219	0.000
H4	WQ → TRS	0.236	3.728	0.000
H5	RA → RQ	0.284	3.161	0.002
H6	RN → RQ	0.471	7.637	0.000
H7	RD → RQ	0.171	2.651	0.008
H8	RQ → TRS	0.302	4.694	0.000
H9	EX → TR	0.819	25.643	0.000
H10	TR → TRS	0.401	5.546	0.000
H11	TRS → IN	0.616	10.968	0.000

Table 4
Evaluation of research hypotheses (Lazada).

Hyp.	Link	β	t-value	p-value
H1	US → WQ	0.562	7.544	0.000
H2	IQ → WQ	0.106	1.992	0.047
H3	SI → WQ	0.244	3.129	0.002
H4	WQ → TRS	0.189	2.628	0.009
H5	RA → RQ	0.257	3.842	0.000
H6	RN → RQ	0.350	6.195	0.000
H7	RD → RQ	0.319	5.311	0.000
H8	RQ → TRS	0.345	3.387	0.001
H9	EX → TR	0.702	15.979	0.000
H10	TR → TRS	0.445	5.287	0.000
H11	TRS → IN	0.590	9.541	0.000

Table 5
R-squares of dependent variables (Amazon).

Dep. variables	Notation	R-Square
Purchase Intention	IN	0.379
Trust	TRS	0.634
Web site Quality	WQ	0.524
Recommendation Quality	RQ	0.583
Transparency	TR	0.671

Table 6
R-Squares of Dependent Variables (Lazada).

Dep. variables	Notation	R-Square
Purchase Intention	IN	0.348
Trust	TRS	0.793
Website Quality	WQ	0.668
Recommendation Quality	RQ	0.608
Transparency	TR	0.492

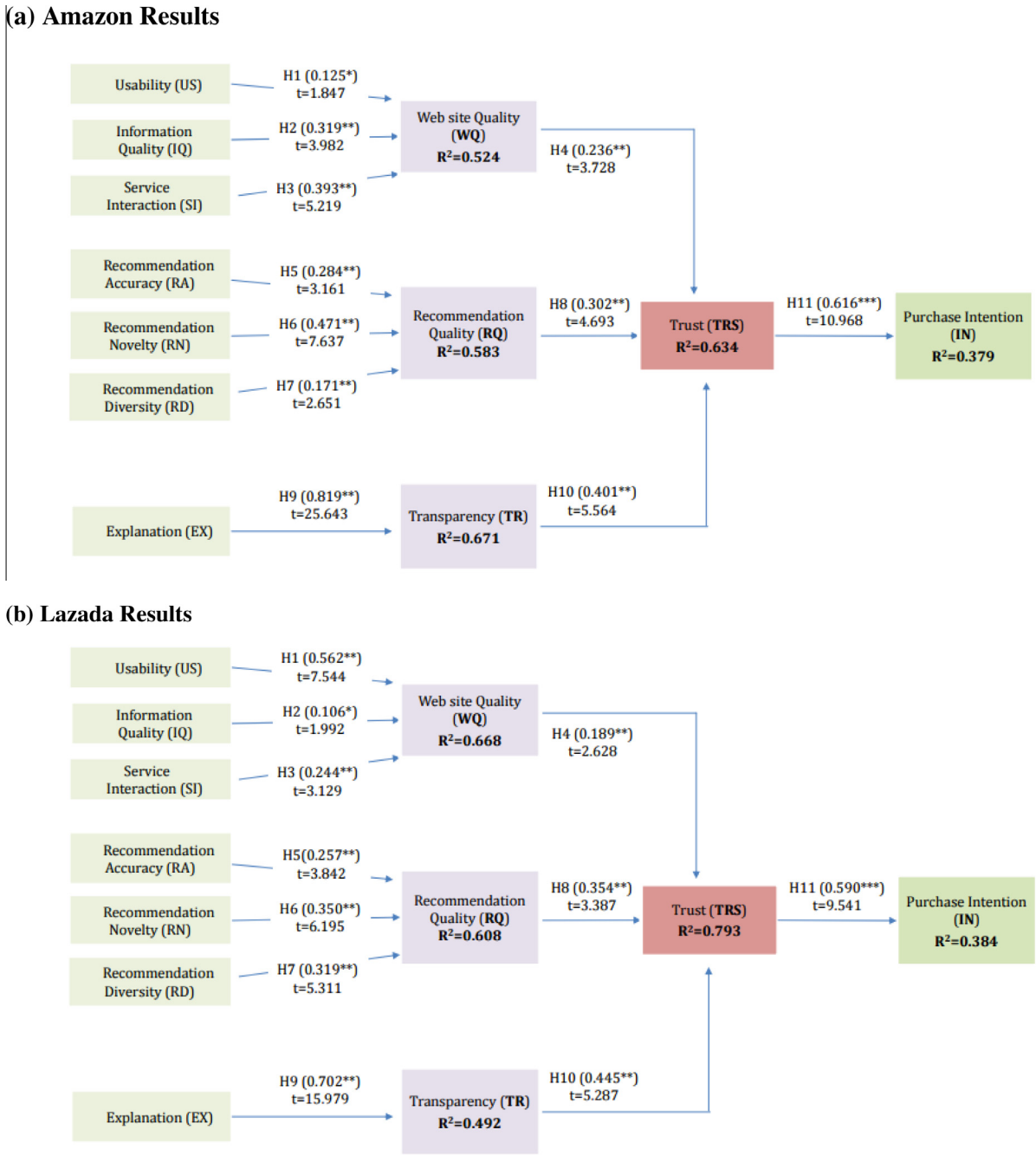

Fig. 2. Summary of results for model fitted with PLS.

- Recommendation accuracy, diversity, and novelty are relevant quality factors in both domains, but with different levels of importance.
- Differences across the two sites also exist with respect to the relative importance of the different factors that impact website quality.

6. Conclusion

6.1. Theoretical contributions

Our work continues existing lines of research that aim to better understand the factors that influence the adoption of e-commerce recommendation agents. We focus on the role of trust, a central concept for the success of recommendation agents (Wang and Benbasat, 2008; Xiao and Benbasat, 2007) and also for e-commerce adoption (Grabner-Kräuter and Kaluscha, 2003; Lim et al., 2006).

We take into account that recommendation agents are almost never stand-alone applications but are usually one of several components of an e-commerce website. We propose a new model that combines aspects of website quality with other trust-building factors. Overall, our survey-based study not only confirmed our Website Quality → Recommendation Trust Hypothesis (H4) that website quality significantly affects consumer trust in an embedded recommendation agent. It also shows the relative importance of the different related factors. An interesting

insight is that the importance of the three main trust-building factors considered in our study is very similar across the two e-commerce websites and product domains – experience versus search products.

When we look at the details of the effects on these main factors, we observe differences across the domains. While the effect of recommendation accuracy on recommendation quality is similar for books and cameras, providing more diverse recommendations led to a higher perceived recommendation quality for digital cameras. Novelty was slightly more appreciated by the participants in the domain of books. This seems plausible: many users of a book recommendation agent may be interested in finding something new, whereas buyers of digital cameras may appreciate a recommendation agent capable of proposing a diverse choice set that does not contain too similar items. Generally, however, the findings corroborate the assumption that being accurate is not enough for building a successful recommendation system (McNee et al., 2006), and other factors such as novelty or diversity can play an important role.

In the context of quality factors for recommendations, familiarity is considered to be trust-building (Xiao and Benbasat, 2007). Users may be leaning to trust a recommendation agent or feel positive about it, if it, among other things, recommends items to them that they already know (Jannach et al., 2015). Assessing the relative importance of item familiarity as a trust-building factor is part of our future work. We did not include it in our present study. Participants did not log in with their usual accounts, and the system could not intentionally recommend items that the user had seen in the past.

Looking at the factors that contribute to website quality, we observe differences. For Lazada.com, the dominant contributing factor is usability, whereas on Amazon.com the contributions of information quality and service interaction were much more pronounced. As a result, since website quality had an impact on trust in both cases, we conclude that all of the factors that we considered can be relevant, but their relative importance seems to depend on the domain or specific application.

Finally, given the strong effect of trust on user purchase intentions, our work further demonstrates the importance of trust as a means to influence the behavioral intentions of online customers.

6.2. Practical implications

From a practical perspective, the dominating role of perceived transparency for the development of trust implies that e-commerce websites in the future should focus more on incorporating explanation mechanisms into their recommendation agents. Concentrating solely on the optimization of machine learning models for accurate recommendations may be insufficient. This is important: algorithms optimized for accuracy in an offline process with historical data are not necessarily the most successful ones in practice (Gomez-Uribe and Hunt, 2015; Garcin et al., 2014).

Amazon.com has implemented advanced forms of explanation for its recommendations to users and puts them into control. Compared to the Lazada.com, the effect of transparency on trust is, possibly as a result, higher on Amazon.com (0.819 versus 0.702). Nonetheless, the effect on Lazada.com also is quite strong though the explanations are limited to labels of the recommendation lists or aggregated community ratings. The underlying recommendation mechanism for the camera domain seems to recommend similar and popular items of the same category. This makes it obvious for users to see how the recommendation list was created. In case of experience products like books, more diverse or even serendipitous recommendations may be helpful for users, but they may necessitate better explanations for the user.

One main practical challenge of providing such explanations lies in the design of the explanation interface. They must be designed so they are easy to comprehend by a heterogeneous user group. Understandability may not be enough though: different types of interfaces can also lead to desired or undesired effects on users (Gedikli et al., 2014).

The observations related to novelty and diversity suggest that the "right" level of these factors depends on the domain and purpose of the recommendation system. If the goal is to point users to novel items and help them discover the item space (e.g., in online music platforms), more diversity and novelty seems appropriate. If, on the other hand, the goal of the recommender is to provide the user with a limited choice set of similar items in a decision-making scenario, a collection of comparable items (i.e., less diversity) may be more valuable for the user. Alternatively, a recommender may recommend substitutes for a currently-viewed item, or recommend complements and accessories (Diehl et al., 2015). One practical implication of these insights is that e-commerce providers should have the intended utility of the recommendation agent and the domain in mind when designing its logic.

6.3. Research limitations

The 150 participants of our study all had a certain background and education level.[3] While students of computer science and IS have many characteristics of typical online shoppers, there may be aspects of a recommendation agent that are more relevant for students than the general user population. Nonetheless, no specific IT knowledge was required for the participants to answer the questionnaire items. And the participants had no particular background in the field of recommender systems, which could have led to a heightened awareness with respect to recommendation quality.

Most participants had several years of experience with both websites. Positive and negative prior experiences at the level of individual participants existed, but were not captured by our research model. The effects of these prior experiences should average out across the participant population. Another possible limitation of using two popular websites is that the relative importance of the quality factors may have been different had we used less well-known websites. General website quality, for example, may be more important for users who have not interacted with the site before. We held the popularity level of the websites constant, so as to not add another salient dimension. The influence of other characteristics on the relative importance of trust-building quality factors is left for the future.

Overall, our observations regarding perceived recommendation quality are based on surveys for two different e-commerce websites and two application domains: books and cameras. Other domains will no doubt have recommendation quality factors that are different in their importance. Finally, even though we used established questionnaire items for the measurement of intention-to-purchase, the demand scenario for the participants was remains was artificial: the participants did not have to make a purchase decision.

[3] The sample size of 150 participants is sufficient for our study according to the heuristic proposed by Hair et al. (2013) for exploratory PLS-SEM models and Boomsma's (1982) rule. Confirmatory SEM models with the same number of variables, indicators, and paths may require larger sample sizes, depending on the minimum effect size (Westland, 2010).

Acknowledgements

The first author would like to thank the Research Management Center (RMC) at Universiti Teknologi Malaysia (UTM) for supporting and funding this research under the Post-Doctoral Fellowship Scheme Grant of vote number Q.J130000.21A2.03E26.

Appendix A. Detailed statistical results and questionnaire items

The section provides the sources of the questionnaire items used in this study. We also include detailed statistical results for both e-commerce platforms (see Table A.11):

- Outer loadings of the reflective constructs and the discriminant validity results based on the cross-loadings matrix (Tables A.1 and A.2)
- Results regarding the internal consistency of the reflective constructs (Tables A.3 and A.4)
- Results of the assessment of convergent validity analysis of the reflective constructs (Table A.5)
- Discriminant validity assessment based on Fornell–Larcker criterion (Tables A.6 and A.7)
- Variance inflation factor (VIF) analysis for the formative construct (Table A.8)
- Significance assessment of the formative construct information quality (Tables A.9 and A.10).

Abbreviations: EX = Explanation; IQ = Information Quality; RA = Recommendation Accuracy; RD = Recommendation Diversity; RN = Recommendation Novelty; RQ = Recommendation Quality; SI = Service Interaction; TR = Transparency; TRS = Trust; US = Usability; WQ = Website Quality; IN = Intention to Purchase.

Table A.1
Outer loadings of the reflective constructs and discriminant validity based on cross-loadings matrix (Amazon).

	EX	IN	RA	RD	RN	RQ	SI	TR	TRS	US	WQ
EX1	0.868	0.401	0.343	0.227	0.306	0.428	0.401	0.670	0.414	0.300	0.377
EX2	0.890	0.479	0.313	0.289	0.357	0.489	0.437	0.745	0.545	0.390	0.478
EX3	0.882	0.389	0.320	0.323	0.331	0.481	0.401	0.743	0.495	0.364	0.398
IN1	0.382	0.833	0.391	0.356	0.436	0.526	0.374	0.455	0.516	0.353	0.450
IN2	0.363	0.842	0.389	0.431	0.534	0.593	0.344	0.485	0.489	0.269	0.444
IN3	0.418	0.737	0.394	0.418	0.310	0.451	0.397	0.479	0.481	0.319	0.476
RA1	0.341	0.479	0.844	0.304	0.507	0.553	0.163	0.443	0.367	0.125	0.244
RA2	0.280	0.355	0.850	0.237	0.373	0.448	0.183	0.310	0.224	0.155	0.149
RA3	0.305	0.382	0.833	0.327	0.484	0.502	0.207	0.358	0.301	0.131	0.158
RD1	0.284	0.459	0.290	0.930	0.392	0.446	0.143	0.292	0.233	0.220	0.306
RD2	0.311	0.467	0.354	0.931	0.446	0.449	0.115	0.343	0.264	0.182	0.322
RN1	0.321	0.497	0.436	0.399	0.881	0.619	0.203	0.375	0.396	0.112	0.250
RN2	0.273	0.369	0.507	0.399	0.846	0.574	0.167	0.358	0.327	0.132	0.182
RN3	0.389	0.517	0.487	0.382	0.891	0.643	0.239	0.443	0.438	0.104	0.293
RQ1	0.451	0.585	0.516	0.501	0.620	0.887	0.350	0.552	0.570	0.221	0.447
RQ2	0.440	0.520	0.480	0.427	0.681	0.870	0.347	0.478	0.501	0.170	0.314
RQ3	0.445	0.530	0.513	0.268	0.455	0.751	0.294	0.541	0.595	0.326	0.400
SI1	0.363	0.382	0.193	0.154	0.305	0.371	0.704	0.370	0.297	0.420	0.400
SI2	0.397	0.372	0.298	0.081	0.203	0.347	0.770	0.359	0.297	0.457	0.473
SI3	0.410	0.459	0.201	0.147	0.236	0.417	0.791	0.425	0.415	0.554	0.623
SI4	0.309	0.344	0.106	0.084	0.099	0.276	0.806	0.275	0.255	0.477	0.542
SI5	0.383	0.294	0.106	0.055	0.149	0.245	0.828	0.263	0.233	0.488	0.504
SI6	0.360	0.302	0.087	0.088	0.117	0.209	0.794	0.246	0.233	0.417	0.482
SI7	0.360	0.367	0.212	0.152	0.196	0.297	0.792	0.419	0.379	0.513	0.530
TR1	0.754	0.476	0.374	0.285	0.430	0.534	0.332	0.863	0.607	0.299	0.426
TR2	0.696	0.552	0.338	0.388	0.379	0.561	0.398	0.887	0.677	0.378	0.579
TR3	0.683	0.501	0.453	0.211	0.362	0.530	0.394	0.854	0.604	0.368	0.498
TRS1	0.421	0.509	0.276	0.124	0.362	0.514	0.304	0.642	0.832	0.258	0.535
TRS2	0.462	0.427	0.291	0.249	0.353	0.548	0.374	0.542	0.767	0.210	0.478
TRS3	0.436	0.525	0.288	0.270	0.342	0.513	0.250	0.536	0.776	0.223	0.427
US1	0.235	0.292	0.073	0.125	0.039	0.200	0.415	0.247	0.167	0.751	0.374
US2	0.275	0.254	0.199	0.190	0.088	0.202	0.446	0.299	0.211	0.729	0.353
US3	0.296	0.250	0.122	0.125	0.034	0.151	0.471	0.276	0.161	0.753	0.354
US4	0.399	0.314	0.117	0.069	0.111	0.213	0.515	0.391	0.296	0.792	0.471
US5	0.310	0.339	0.160	0.237	0.201	0.255	0.461	0.304	0.189	0.788	0.459
US6	0.257	0.308	0.110	0.166	0.085	0.290	0.535	0.263	0.283	0.752	0.472
US7	0.352	0.271	0.103	0.238	0.165	0.193	0.366	0.339	0.223	0.695	0.400
US8	0.275	0.309	0.095	0.155	0.046	0.168	0.449	0.283	0.196	0.764	0.387
WQ1	0.437	0.453	0.173	0.324	0.184	0.367	0.574	0.500	0.517	0.568	0.847
WQ2	0.425	0.516	0.159	0.245	0.226	0.428	0.611	0.503	0.554	0.494	0.903
WQ3	0.404	0.539	0.262	0.331	0.337	0.434	0.557	0.533	0.544	0.395	0.909

The grey shaded areas show that the loading of each measurement item on its assigned construct is larger than its loading on any other construct in the model.

Table A.2
Outer loadings of the reflective constructs and discriminant validity based on cross-loadings matrix (Lazada).

	EX	IN	RA	RD	RN	RQ	SI	TR	TRS	US	WQ
EX1	0.901	0.399	0.379	0.465	0.442	0.633	0.572	0.612	0.636	0.596	0.625
EX2	0.898	0.333	0.458	0.466	0.438	0.705	0.482	0.588	0.609	0.508	0.596
EX3	0.936	0.431	0.417	0.545	0.481	0.748	0.605	0.709	0.737	0.633	0.690
IN1	0.364	0.848	0.153	0.167	0.417	0.427	0.294	0.349	0.512	0.232	0.330
IN2	0.374	0.844	0.225	0.305	0.428	0.436	0.256	0.344	0.464	0.233	0.320
IN3	0.345	0.840	0.127	0.275	0.402	0.386	0.288	0.407	0.514	0.253	0.296
RA1	0.365	0.250	0.758	0.417	0.451	0.472	0.285	0.362	0.411	0.313	0.333
RA2	0.305	0.073	0.852	0.420	0.435	0.497	0.225	0.254	0.281	0.246	0.247
RA3	0.449	0.169	0.859	0.515	0.530	0.588	0.389	0.434	0.439	0.445	0.432
RD1	0.481	0.272	0.521	0.934	0.525	0.601	0.393	0.542	0.543	0.439	0.538
RD2	0.533	0.276	0.510	0.939	0.508	0.623	0.413	0.574	0.564	0.406	0.553
RN1	0.391	0.448	0.525	0.434	0.878	0.590	0.307	0.404	0.489	0.325	0.379
RN2	0.477	0.418	0.499	0.522	0.904	0.595	0.350	0.502	0.517	0.336	0.410
RN3	0.469	0.454	0.522	0.523	0.902	0.623	0.390	0.491	0.542	0.365	0.415
RQ1	0.695	0.454	0.566	0.591	0.611	0.888	0.511	0.713	0.762	0.565	0.553
RQ2	0.703	0.399	0.605	0.600	0.607	0.933	0.492	0.731	0.742	0.567	0.597
RQ3	0.677	0.486	0.550	0.583	0.611	0.893	0.492	0.708	0.720	0.519	0.601
SI1	0.522	0.271	0.319	0.411	0.400	0.452	0.849	0.454	0.441	0.675	0.635
SI2	0.476	0.311	0.328	0.395	0.327	0.477	0.851	0.426	0.436	0.681	0.638
SI3	0.531	0.323	0.366	0.353	0.384	0.502	0.818	0.428	0.468	0.662	0.641
SI4	0.554	0.236	0.319	0.305	0.306	0.477	0.849	0.416	0.460	0.743	0.628
SI5	0.563	0.258	0.333	0.390	0.369	0.504	0.871	0.424	0.450	0.750	0.624
SI6	0.482	0.272	0.226	0.297	0.237	0.420	0.836	0.386	0.405	0.664	0.578
SI7	0.471	0.291	0.282	0.395	0.278	0.419	0.844	0.403	0.418	0.625	0.571
TR1	0.586	0.392	0.302	0.444	0.447	0.606	0.432	0.809	0.649	0.477	0.509
TR2	0.621	0.358	0.373	0.547	0.462	0.706	0.427	0.890	0.737	0.553	0.603
TR3	0.587	0.365	0.417	0.528	0.425	0.710	0.413	0.855	0.770	0.504	0.591
TRS1	0.661	0.506	0.386	0.503	0.513	0.711	0.430	0.742	0.876	0.499	0.569
TRS2	0.625	0.451	0.394	0.486	0.425	0.711	0.407	0.741	0.861	0.532	0.644
TRS3	0.566	0.544	0.390	0.515	0.530	0.663	0.489	0.666	0.806	0.461	0.591
US1	0.586	0.235	0.371	0.372	0.338	0.533	0.692	0.498	0.495	0.832	0.671
US2	0.568	0.165	0.337	0.383	0.277	0.516	0.721	0.509	0.466	0.907	0.692
US3	0.534	0.232	0.292	0.381	0.250	0.482	0.706	0.509	0.490	0.894	0.693
US4	0.489	0.219	0.387	0.383	0.325	0.530	0.685	0.529	0.498	0.871	0.659
US5	0.608	0.341	0.371	0.413	0.421	0.577	0.729	0.583	0.542	0.871	0.720
US6	0.500	0.275	0.393	0.352	0.346	0.539	0.731	0.501	0.522	0.901	0.697
US7	0.595	0.326	0.388	0.453	0.414	0.554	0.700	0.548	0.566	0.836	0.739
US8	0.580	0.186	0.353	0.414	0.300	0.529	0.719	0.526	0.518	0.890	0.716
WQ1	0.580	0.294	0.386	0.515	0.396	0.546	0.655	0.597	0.588	0.732	0.904
WQ2	0.658	0.269	0.349	0.492	0.406	0.584	0.670	0.629	0.622	0.745	0.911
WQ3	0.633	0.435	0.375	0.549	0.396	0.591	0.626	0.557	0.683	0.658	0.855

The grey shaded areas show that the loading of each measurement item on its assigned construct is larger than its loading on any other construct in the model.

Table A.3
Internal consistency of the reflective constructs (Amazon).

Construct	Composite reliability	Cronbach's Alpha
Explanation	0.911	0.854
Purchase Intention	0.846	0.726
Recommendation Accuracy	0.880	0.796
Recommendation Diversity	0.928	0.846
Recommendation Novelty	0.906	0.785
Recommendation Quality	0.876	0.785
Service Interaction	0.918	0.896
Transparency	0.902	0.836
Trust	0.834	0.702
Usability	0.913	0.891
Website Quality	0.917	0.864

Table A.4
Internal consistency of the reflective constructs (Lazada).

Construct	Composite reliability	Cronbach's Alpha
Explanation	0.937	0.899
Intention	0.881	0.798
Recommendation Accuracy	0.864	0.763
Recommendation Diversity	0.934	0.859
Recommendation Novelty	0.923	0.875
Recommendation Quality	0.931	0.889
Service Interaction	0.946	0.933
Transparency	0.888	0.811
Trust	0.885	0.804
Usability	0.963	0.956
Website Quality	0.920	0.869

Table A.5
Convergent validity of the reflective constructs.

Construct (Amazon)	AVE	Construct (Lazada)	AVE
Explanation	0.774	Explanation	0.831
Purchase Intention	0.648	Purchase Intention	0.712
Recommendation Accuracy	0.709	Recommendation Accuracy	0.679
Recommendation Diversity	0.866	Recommendation Diversity	0.877
Recommendation Novelty	0.762	Recommendation Novelty	0.801
Recommendation Quality	0.703	Recommendation Quality	0.819
Service Interaction	0.615	Service Interaction	0.715
Transparency	0.754	Transparency	0.726
Trust	0.627	Trust	0.720
Usability	0.568	Usability	0.766
Website Quality	0.786	Website Quality	0.793

Table A.6
Discriminant validity based on the Fornell–Larcker criterion. Diagonal elements (in bold) represent the square root of the AVE (Amazon).

	EX	IN	IQ	RA	RD	RN	RQ	SI	TR	TRS	US	WQ
EX	**0.880**											
IN	0.481	**0.805**										
IQ	0.420	0.394	**formt**.									
RA	0.369	0.486	0.137	**0.842**								
RD	0.320	0.498	0.254	0.346	**0.931**							
RN	0.377	0.531	0.213	0.545	0.450	**0.873**						
RQ	0.531	0.651	0.363	0.599	0.481	0.702	**0.838**					
SI	0.470	0.461	0.583	0.218	0.139	0.234	0.395	**0.784**				
TR	0.819	0.587	0.413	0.445	0.341	0.450	0.624	0.431	**0.868**			
TRS	0.554	0.616	0.416	0.359	0.267	0.445	0.661	0.389	0.726	**0.792**		
US	0.401	0.391	0.576	0.162	0.216	0.132	0.282	0.610	0.401	0.291	**0.753**	
WQ	0.476	0.567	0.621	0.222	0.338	0.279	0.462	0.656	0.577	0.607	0.549	**0.887**

Table A.7
Discriminant validity based on the Fornell–Larcker criterion. Diagonal elements (in bold) represent the square root of the AVE (Lazada).

	EX	IN	IQ	RA	RD	RN	RQ	SI	TR	TRS	US	WQ
EX	**0.912**											
IN	0.428	**0.844**										
IQ	0.321	0.265	**formt.**									
RA	0.457	0.197	0.201	**0.824**								
RD	0.542	0.293	0.150	0.550	**0.936**							
RN	0.499	0.492	0.201	0.576	0.551	**0.895**						
RQ	0.764	0.493	0.264	0.634	0.654	0.674	**0.905**					
SI	0.609	0.332	0.290	0.369	0.431	0.391	0.551	**0.845**				
TR	0.702	0.435	0.176	0.429	0.596	0.521	0.793	0.497	**0.852**			
TRS	0.728	0.590	0.285	0.460	0.591	0.577	0.820	0.521	0.845	**0.848**		
US	0.638	0.284	0.368	0.413	0.451	0.383	0.609	0.812	0.601	0.586	**0.875**	
WQ	0.701	0.374	0.390	0.415	0.583	0.449	0.645	0.731	0.668	0.709	0.799	**0.891**

Table A.8
Variance inflation factor (VIF) for the formative construct.

Construct	Measurement item	VIF (Amazon)	VIF (Lazada)
Information Quality	IQ1	1.670	1.598
	IQ2	2.217	3.641
	IQ3	1.794	1.825
	IQ4	2.019	2.021
	IQ5	2.569	1.595
	IQ6	2.184	4.059
	IQ7	1.936	1.489

Table A.9
Significance assessment of the formative construct (Amazon).

Construct	Measurement item	Significance			
		Outer weight	Outer loading	t-value$_{>1.96}$	p-value
Information Quality	IQ1	0.211	0.721	1.618	0.106
	IQ2	0.452	0.845	3.271	0.001
	IQ3	0.109	0.697	0.783	0.434
	IQ4	0.221	0.790	1.840	0.066
	IQ5	0.241	0.840	1.704	0.089
	IQ6	-0.114	0.617	0.861	0.390
	IQ7	0.119	0.696	0.956	0.340

Table A.10
Significance assessment of the formative construct (Lazada).

Construct	Measurement item	Significance			
		Outer weight	Outer loading	t-value$_{>1.96}$	p-value
Information Quality	IQ1	0.194	0.672	0.855	0.393
	IQ2	0.470	0.792	1.135	0.257
	IQ3	0.351	0.774	1.482	0.139
	IQ4	0.315	0.779	1.311	0.190
	IQ5	0.079	0.625	0.368	0.713
	IQ6	-0.015	0.780	0.034	0.973
	IQ7	-0.123	0.463	0.528	0.598

Table A.11
References for questionnaire items.

Construct	Source
Usability (US)	Barnes and Vidgen (2002 2003)
Information Quality (IQ)	Barnes and Vidgen (2002, 2003)
Service Interaction (SI)	Barnes and Vidgen (2002, 2003)
Website Quality (WQ)	Al-Qeisi et al. (2014), Chang et al. (2014 McKnight et al. (2002)
Recommendation Accuracy (RA)	Pu et al. (2011)
Recommendation Novelty (RN)	Pu et al. (2011)
Recommendation Diversity (RD)	Pu et al. (2011)
Recommendation Quality (RQ)	Pu et al. (2011), Yoon et al. (2013)
Explanation (EX)	Pu et al. (2011)
Transparency (TR)	Pu et al. (2011)
Trust (TRS)	Kim et al. (2008), Ganguly et al. (2010)
Purchase Intention (IN)	Kim et al. (2008), Ganguly et al. (2010)

References

Adamopoulou, A.A., Symeonidis, A.L., 2014. A simulation testbed for analyzing trust and reputation mechanisms in unreliable online markets. Electron. Commer. Res. Appl. 13 (5), 368–386.
Aksoy, L., Bloom, P.N., Lurie, N.H., Cooil, B., 2006. Should recommendation agents think like people? J. Serv. Res. 8 (4), 297–315.
Al-Qeisi, K., Dennis, C., Alamanos, E., Jayawardhena, C., 2014. Website design quality and usage behavior: unified theory of acceptance and use of technology. J. Business Res. 67 (11), 2282–2290.
Baier, D., Stüber, E., 2010. Acceptance of recommendations to buy in online retailing. J. Retail. Consum. Serv. 17 (3), 173–180.
Barnes, S.J., Vidgen, R.T., 2002. An integrative approach to the assessment of e-commerce quality. J. Electron. Commerce Res. 3 (3), 114–127.
Barnes, S.J., Vidgen, R., 2003. Measuring web site quality improvements: a case study of the forum on strategic management knowledge exchange. Ind. Manage. Data Syst. 103 (5), 297–309.
Bodoff, D., Ho, S.Y., 2015. Effectiveness of website personalization: does the presence of personalized recommendations cannibalize sampling of other items? Int. J. Electron. Commerce 20 (2), 208–235.
Boomsma, H., 1982. Robustness of LISREL against small sample sizes in factor analysis models. In: Joreskog, K., Wold, H. (Eds.), Systems under Indirect Observation: Causality, Structure, Prediction. Elsevier, Amsterdam, Netherlands, pp. 149–173.
Chang, H.H., Chen, S.W., 2008. The impact of online store environment cues on purchase intention: trust and perceived risk as a mediator. Online Inf. Rev. 32 (6), 818–841.
Chang, K.C., Kuo, N.T., Hsu, C.L., Cheng, Y.S., 2014. The impact of website quality and perceived trust on customer purchase intention in the hotel sector: website brand and perceived value as moderators. Int. J. Innov. Manage. Technol. 5 (4), 255–260.
Chau, P.Y.K., Ho, S.Y., Ho, K.K.W., Yao, Y., 2013. Examining the effects of malfunctioning personalized services on online users' distrust and behaviors. Decis. Support Syst. 56, 180–191.
Chopra, K., Wallace, W.A., 2003. Trust in electronic environments. In: Proceedings of the 36th Annual Hawaii International Conference on System Sciences, Vol. 9, IEEE Computer Society Press, Washington, DC.
Corbitt, B.J., Thanasankit, T., Yi, H., 2003. Trust and e-commerce: a study of consumer perceptions. Electron. Commer. Res. Appl. 2 (3), 203–215.
Cremonesi, P., Garzotto, F., Negro, S., Papadopoulos, A., Turrin, R., 2011. Comparative evaluation of recommender system quality. In: Proceedings of CHI '11, Extended Abstracts on Human Factors in Computing Systems. A.C.M. Press, New York, pp. 1927–1932.
de Ruyter, K., Wetzels, M., Kleijnen, M., 2001. Customer adoption of eservice: an experimental study. Int. J. Serv. Ind. Manage. 12 (2), 184–207.
Dias, M.B., Locher, D., Li, M., El-Deredy, W., Lisboa, P.J., 2008. The value of personalised recommender systems to e-business: a case study. In: Proceedings of the 2008 ACM Conference on Recommender Systems. ACM Press, New York, NY, pp. 291–294.
Diehl, K., van Herpen, E., Lamberton, C., 2015. Organizing products with complements versus substitutes: effects on store preferences as a function of effort and assortment perceptions. J. Retail. 91 (1), 1–18.
Ekstrand, M.D., Harper, F.M., Willemsen, M.C., Konstan, J.A., 2014. User perception of differences in recommender algorithms. In: Proceedings of the 2014 ACM Conference on Recommender Systems. ACM Press, New York, NY, pp. 161–168.
Fleder, D., Hosanagar, K., 2009. Blockbuster culture's next rise or fall: the impact of recommender systems on sales diversity. Manage. Sci. 55 (5), 697–712.
Fornell, C., Bookstein, F.L., 1982. Two structural equation models: Lisrel and pls applied to consumer exit-voice theory. J. Mark. Res. 19 (4), 440–452.
Fornell, C., Larcker, D.F., 1981. Evaluating structural equation models with unobservable variables and measurement error. J. Mark. Res. 18 (1), 39–50.
Ganguly, B., Dash, S.B., Cyr, D., Head, M., 2010. The effects of website design on purchase intention in online shopping: the mediating role of trust and the moderating role of culture. Int. J. Electron. Business 8 (4–5), 302–330.
Garcin, F., Faltings, B., Donatsch, O., Alazzawi, A., Bruttin, C., Huber, A., 2014. Offline and online evaluation of news recommender systems at swissinfo.ch. In: Proceedings of the 2014 ACM Conference on Recommender Systems, Foster City, Silicon Valley, CA. ACM Press, New York, NY, pp. 169–176.
Ge, M., Delgado-Battenfeld, C., Jannach, D., 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In: Proceedings of the Fourth ACM Conference on Recommender Systems. ACM Press, New York, NY, pp. 257–260.
Gedikli, F., Jannach, D., Ge, M., 2014. How should I explain? A comparison of different explanation types for recommender systems. Int. J. Hum Comput Stud. 72 (4), 367–382.
Gefen, D., 2000. E-commerce: the role of familiarity and trust. Omega 28 (6), 725–737.
Gefen, D., Karahanna, E., Straub, D., 2003. Inexperience and experience with online stores: the importance of TAM and trust. IEEE Trans. Eng. Manage. 50 (3), 307–321.
Gomez-Uribe, C.A., Hunt, N., 2015. The Netflix recommender system: algorithms, business value, and innovation. ACM Trans. Manage. Inf. Syst. 6 (4), 1–13:19.
Grabner-Kräuter, S., Kaluscha, E., 2003. Empirical research in on-line trust: a review and critical assessment. Int. J. Hum Comput. Stud. 58 (6), 783–812.
Hair, J., Anderson, R., Babin, B., 1998. Multivariate Data Analysis. Prentice Hall, Englewood Cliffs, NJ.
Hair, J.F., Ringle, C.M., Sarstedt, M., 2011. PLS-SEM: indeed a silver bullet. J. Market. Theory Pract. 19 (2), 139–152.
Hair, J., Hult, G., Ringle, C., Sarstedt, M.A., 2013. Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). Sage, Thousand Oaks, CA.
Hampton-Sosa, W., Koufaris, M., 2005. The effect of web site perceptions on initial trust in the owner company. J. Electron. Commerce Res. 10 (1), 55–81.
Hasan, L., Abuelrub, E., 2011. Assessing the quality of web sites. Appl. Comput. Inf. 9 (1), 11–29.
Häubl, G., Murray, K.B., 2006. Double agents: assessing the role of electronic product recommendation systems. MIT Sloan Manage. Rev. 47 (3), 8–12.
Häubl, G., Trifts, V., 2000. Consumer decision making in online shopping environments: the effects of interactive decision aids. Market. Sci. 19 (1), 4–21.
Henseler, J., Ringle, C.M., Sinkovics, R.R., 2009. The use of partial least squares path modeling in international marketing. Ch. 19 in: B. Stottinger, B.B. Schegelmilch, S. Zou (eds.), International Marketing in the Fast Changing World, Advances in International Marketing, 26, Emerald Insight, Bingley, UK, pp. 277–319.
Herlocker, J.L., Konstan, J.A., Riedl, J., 2000. Explaining collaborative filtering recommendations. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work. ACM Press, New York, NY, pp. 241–250.
Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T., 2004. Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. 22 (1), 5–53.
Huang, Z., Benyoucef, M., 2013. From e-commerce to social commerce: a close look at design features. Electron. Commer. Res. Appl. 12 (4), 246–259.
Jannach, D., Hegelich, K., 2009. A case study on the effectiveness of recommendations in the mobile interne. In: Proceedings of the 2009 ACM Conference on Recommender Systems. ACM Press, New York, NY, pp. 205–208.
Jannach, D., Zanker, M., Ge, M., Gröning, M., 2012. Recommender systems in computer science and information systems: a landscape of research. In: Proceedings 13th International Conference on E-Commerce and Web Technologies. Springer, Heidelberg, Germany, pp. 76–87.
Jannach, D., Lerche, L., Jugovac, M., 2015. Item familiarity as a possible confounding factor in user-centric recommender systems evaluation. I-com J. Interact. Media 14 (1), 29–40.
Jarvenpaa, S.L., Tractinsky, N., Saarinen, L., 1999. Consumer trust in an Internet store: a cross-cultural validation. J. Comput. Mediated Commun. 5 (2).
Jiang, J.J., Klein, G., Carr, C.L., 2002. Measuring information system service quality: SERVQUAL from the other side. MIS Q. 26 (2), 145–166.
Johnson, H., Johnson, P., 1993. Explanation facilities and interactive systems. In: Proceedings of the 1st International Conference on Intelligent User Interfaces. ACM Press, New York, NY, pp. 159–166.
Jöreskog, K., Sörbom, D., 1993. LISREL 8: Structural Equation Modeling With the SIMPLIS Command Language. Scientific Software International, Hillsdale, NJ.
Karimov, F., Brengman, M., Van Hove, L., 2011. The effect of website design dimensions on initial trust: a synthesis of the empirical literature. J. Electron. Commerce Res. 12 (4), 272–301.

Kijsanayotin, B., Pannarunothai, S., Speedie, S.M., 2009. Factors influencing health information technology adoption in Thailand's community health centers: applying the UTAUT model. Int. J. Med. Informatics 78 (6), 404–416.
Kim, D.J., Ferrin, D.L., Rao, H.R., 2008. A trust-based consumer decision-making model in electronic commerce: the role of trust, perceived risk, and their antecedents. Decis. Support Syst. 44 (2), 544–564.
Kim, C., Galliers, R.D., Shin, N., Ryoo, J.H., Kim, J., 2012. Factors influencing internet shopping value and customer repurchase intention. Electron. Commer. Res. Appl. 11 (4), 374–387.
Knijnenburg, B., Willemsen, M., Gantner, Z., Soncu, H., Newell, C., 2012. Explaining the user experience of recommender systems. User Model. User-Adap. Inter. 22 (4), 441–504.
Komiak, S.Y.X., Benbasat, I., 2006. The effects of personalization and familiarity on trust and adoption of recommendation agents. MIS Q. 30 (4), 941–960.
Konstan, J.A., Riedl, J., 2012. Recommender systems: from algorithms to user experience. User Model. User-Adap. Inter. 22 (1–2), 101–123.
Koufaris, M., Hampton-Sosa, W., 2002. Customer trust online: examining the role of the experience with the web-site. In: Department of Statistics and Computer Information Systems Working Paper Series. Zicklin School of Business, Baruch College, NY.
Ku, Y.C., Tai, Y.M., 2013. What happens when recommendation system meets reputation system? The impact of recommendation information on purchase intention. In: Proceedings of the 46th Hawaii International Conference on System Sciences. IEEE Computer Society Press, Washington, DC.
Lerche, L., Jannach, D., Ludewig, M., 2016. On the value of reminders within e-commerce recommendations. In: Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization, pp. 27–35.
Liang, T.P., Lai, H.J., Ku, Y.C., 2007. Personalized content recommendation and user satisfaction: theoretical synthesis and empirical findings. J. Manage. Inf. Syst. 23 (3), 45–70.
Lim, K., Sia, C., Lee, M., Benbasat, I., 2006. Do I trust you online, and if so, will I buy? An empirical study of two trust-building strategies. J. Manage. Inf. Syst. 23 (2), 233–266.
Linden, G., Smith, B., York, J., 2003. Amazon.com recommendations: item-to-item collaborative filtering. IEEE Internet Comput. 7 (1), 76–80.
Mayer, R.C., Davis, J.H., Schoorman, F.D., 1995. An integrative model of organizational trust. Acad. Manag. Rev. 20 (3), 709–734.
McKnight, D.H., Cummings, L.L., Chervany, N.L., 1998. Initial trust formation in new organizational relationships. Acad. Manag. Rev. 23 (3), 473–490.
McKnight, D.H., Choudhury, V., Kacmar, C., 2002. The impact of initial consumer trust on intentions to transact with a web site: a trust building model. J. Strategic Inf. Syst. 11 (3), 197–323.
McNee, S. M., Riedl, J., Konstan, J. A., 2006. Being accurate is not enough: How accuracy metrics have hurt recommender systems. In: Proceedings of the 2006 Conference on Human Factors in Computing Systems, Montreal, Canada, ACM Press, New York, NY, pp. 1097–1101.
Moody, G.D., Galletta, D.F., Lowry, P.B., 2014. When trust and distrust collide online: the engenderment and role of consumer ambivalence in online consumer behavior. Electron. Commer. Res. Appl. 13 (4), 266–282.
Nunnally, J.C., 1978. Psychometric Theory. McGraw-Hill, New York.
Olsina, L., Rossi, G., 2002. Measuring web application quality with WebQEM. IEEE Multimedia 9 (4), 20–29.
Pavlou, P.A., Chellappa, R.K., 2001. The role of perceived privacy and perceived security in the development of trust in electronic commerce transactions. Inf. Syst. Res. 11, 18–36.
Polatidis, N., Georgiadis, C.K., 2014. Factors influencing the quality of the user experience in ubiquitous recommender systems. In: Proceedings of the 2nd International Conference on Distributed, Ambient, and Pervasive Interactions, Crete, Greece, June 2014. Springer, Heidelberg, Germany, pp. 369–379.
Pu, P., Chen, L., 2006. Trust building with explanation interfaces. In: Proceedings of the 11th International Conference on Intelligent User Interfaces, pp. 93–100.
Pu, P., Chen, L., 2007. Trust-inspiring explanation interfaces for recommender systems. Knowl.-Based Syst. 20 (6), 542–556.
Pu, P., Chen, L., Hu, R., 2011. A user-centric evaluation framework for recommender systems. In: Proceedings of the 2011 ACM Conference on Recommender Systems, ACM Press, New York, NY, pp. 157–164.
Ringle, C. M., Sarstedt, M., Straub, D., 2012. A critical look at the use of PLS-SEM. MIS Q. 36 (1), iii–xiv.
Said, A., Fields, B., Jain, B. J., Albayrak, S., 2013. User-centric evaluation of a k-furthest neighbor collaborative filtering recommender algorithm. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, ACM Press, New York, NY, pp. 1399–1408.
Sekaran, U., Bougie, R., 2010. Research Methods for Business: A Skill Building Approach. John Wiley, New York, NY.
Sinha, R. R., Swearingen, K., 2001. Comparing recommendations made by online systems and friends. In: DELOS-NSF Workshop on Personalisation and Recommender Systems in Digital Libraries, ACM Press, New York, NY.
Sinha, R., Swearingen, K., 2002. The role of transparency in recommender systems. In: CHI '02, Extended Abstracts on Human Factors in Computing Systems, ACM Press, New York, NY, pp. 830–831.
Sun, Y., Chong, W.K., Han, Y.S., Rho, S., Lok, K., 2015a. Key factors affecting user experience of mobile recommendation systems. In: G.C. Yang, S.I. Ao, X. Huang, O. Castillo (eds.), Transactions on Engineering Technologies, International MultiConference of Engineers and Computer Scientists, Springer, Heidelberg, Germany.
Sun, Y., Chong, W. K., Man, K.L., Rho, S., Xie, D., 2015b. Exploring critical success factors of mobile recommendation systems: The end user perspective. In: G.C. Yang, S.I. Ao, X. Huang, O. Castillo (eds.), Transactions on Engineering Technologies, International MultiConference of Engineers and Computer Scientists, Springer, Heidelberg, German.
Swearingen, K., Sinha, R., 2002. Interaction design for recommender systems. Design. Interact. Syst. 6 (12), 312–334.
Tintarev, N., Masthoff, J., 2007a. Effective explanations of recommendations: User-centered design. In: Proceedings of the 2007 ACM Conference on Recommender Systems, ACM Press, New York, NY, pp. 153–156.
Tintarev, N., Masthoff, J., 2007b. A survey of explanations in recommender systems. In: Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop. IEEE Computer Society Press, Los Alamitos, CA, pp. 801–810.
Urbach, N., Ahlemann, F., 2010. Structural equation modeling in information systems research using partial least squares. J. Inf. Technol. Theory Appl. 11 (2), 5–40.
Vargas, S., Castells, P., 2011. Rank and relevance in novelty and diversity metrics for recommender systems. In: Chicago, I.L. (Ed.), Proceedings of the 2011 ACM Conference on Recommender Systems. ACM Press, New York, NY, pp. 109–116.
Wang, W., Benbasat, I., 2005. Trust in and adoption of online recommendation agents. J. Assoc. Inf. Syst. 6 (3), 72–100.
Wang, W., Benbasat, I., 2008. Attributions of trust in decision support technologies: a study of recommendation agents for e-commerce. J. Manage. Inf. Syst. 24 (4), 249–273.
Wang, H., Xie, M., Goh, T.N., 1999. Service quality of Internet search engines. J. Inf. Sci. 25 (6), 499–507.
Wang, W., Qiu, L., Kim, D., Benbasat, I., 2016. Effects of rational and social appeals of online recommendation agents on cognition- and affect-based trust. Decis. Support Syst. 86, 48–60.
Westland, J.C., 2010. Lower bounds on sample size in structural equation modeling. Electron. Commer. Res. Appl. 9 (6), 476–487.
Wilcox, J.B., Howell, R.D., Breivik, E., 2008. Questions about formative measurement. J. Business Res. 61 (12), 1219–1228.
Xiao, B., Benbasat, I., 2007. E-commerce product recommendation agents: use, characteristics, and impact. MIS Q. 31 (1), 137–209.
Yan, Z., Holtmanns, S., 2007. Trust modeling and management: from social trust to digital trust. In: Subramanian, R. (Ed.), Computer Security, Privacy and Politics: Current Issues. Challenges and Solutions, IGI Global, IRM Press, Hershey, PA, pp. 290–323.
Yoon, V.Y., Hostler, R.E., Guo, Z., Guimaraes, T., 2013. Assessing the moderating effect of consumer product knowledge and online shopping: experience on using recommendation agents for customer loyalty. Decis. Support Syst. 55 (4), 883–893.
Zanker, M., 2012. The influence of knowledgeable explanations on users' perception of a recommender system. In: Proceedings of the Sixth ACM Conference on Recommender Systems, ACM Press, New York, NY, pp. 269–272.
Zanker, M., Bricman, M., Gordea, S., Jannach, D., Jessenitschnig, M., 2006. Persuasive online-selling in quality & taste domains. In: Proceedings of the 7th International Conference on Electronic Commerce and Web Technologies. Springer, Heidelberg, Germnay, pp. 51–60.
Ziegler, C.N., McNee, S.M., Konstan, J.A., Lausen, G., 2005. Improving recommendation lists through topic diversification. In: Proceedings of the 14th International Conference on World Wide Web, Chiba, Japan, ACM Press, New York, NY, pp. 22–32.
Zimmerman, J., Kurapati, K., 2002. Exposing profiles to build trust in a recommender. In: CHI' 02, Extended Abstracts on Human Factors in Computing Systems. ACM Press, New York, NY, pp. 608–609.