Privacy Concerns in Online Recommender Systems: Influences of Control and User Data Input

Bo Zhang
College of Communications
Pennsylvania State University &
Samsung Research America
buz114@psu.edu

Na Wang
College of Info Sciences & Technology
Pennsylvania State University &
Samsung Research America
nzw109@ist.psu.edu

Hongxia Jin
Samsung Research America
75 West Plumeria Dr.
San Jose, CA 95134
hongxia.jin@samsung.com

ABSTRACT

Recommender systems (e.g., Amazon.com) provide users with tailored products and services, which have the potential to induce user privacy concerns. Although system designers have been actively developing algorithms to introduce user control mechanisms, it remains unclear whether such control is effective in alleviating privacy concerns. It also is unclear how data type affects this relationship. To determine the psychological mechanisms of user privacy concerns in a recommender system, we conducted a scenario-based online experiment (*N* = 385). Users' privacy concerns were measured in relation to different data input (explicit vs. implicit) and control (present vs. absent) scenarios. Results show that a control mechanism can effectively reduce users' concerns over implicit user data input (i.e., purchase history) but not over explicit user data input (i.e., product ratings). We also demonstrate that control can influence privacy concerns via users' perceived value of disclosure. These findings question the effectiveness of user control mechanisms in recommender systems with explicit data input. Additionally, our item categorization provides a reference for future personalized recommendations and future analyses.

Categories and Subject Descriptors

Recommender

Keywords

Recommender system, Privacy, User Control, User Data Input, Privacy Concern, Information Disclosure

1. INTRODUCTION

Online recommender systems (e.g., Amazon.com, Yelp) have become unprecedentedly popular with the advancement of information tracking and prediction algorithms. Tracing extensive data about user preferences and behaviors, recommender systems can help users make better and faster choices specifically tailored for them in multiple areas of their lives (e.g., e-commerce purchasing, movie viewing, restaurant picking) [50]. This not only reduces users' cognitive load, but also provides them with more relevant and valuable services and products. Striving for more accurate predictions, a vast body of research has been devoted to creating and refining algorithms on recommender platforms [8, 30].

However, these personalized recommendations also pose severe threats to online users' privacy. To accurately predict what users want and need, recommender systems usually rely on a large amount of user data collected out of users' expectations [32], thereby inducing privacy concerns [3, 46]. The concerns, in return, affect users' evaluations of the system [29]. This user data includes demographic information that can point to one's unique identity (e.g., email addresses and social security numbers), as well as product-related footprints users leave online through web browsing and purchasing, hinting at one's tastes and habits. Due to the variation in sensitivity among numerous pieces of user data, it is inefficient to implement a holistic protection mechanism at the cost of recommendation quality. Hence, it becomes imperative to differentiate sensitive information from non-sensitive information and determine users' concerns about them in a recommender context respectively; that way, system developers can create suitable remedies for balancing prediction quality and privacy loss.

In addition to various types of user data, the channels they are collected through—either explicit (e.g., product rating) or implicit (e.g., purchase history)—also bring about privacy concerns [6]. Both approaches are meant to offer service providers extant data for predicting user needs. Explicit data input puts users in a conscious situation and requires their effort to complete the process, whereas implicit input is processed automatically, usually without user awareness. The former may empower users with a sense of control but make the privacy issue more salient, whereas the latter may provide users with more seamless convenience but also come with a sense of intrusiveness that leads to privacy concerns. This study examines how these two types of data input affect users' privacy concerns.

In addressing privacy concern issues in recommender systems, much attention has been put on creating solutions, such as granting users control over information release [31] or providing disclosure justifications [27]. In principle, control enables users to better manage their information flow and make decisions on information sharing, so as to reduce concerns about privacy. In reality, active user control could increase users' cognitive load, which may impede the expected effectiveness. Also, it is unclear whether the presence of a control mechanism will moderate the effect of data input on privacy concerns.

Symposium on Usable Privacy and Security (SOUPS) 2014, July 9-11, 2014, Menlo Park, CA.

In addition to investigating the effects of data input and user control on privacy concerns, this study also probes the underlying psychological mechanisms that could explain the causes of privacy concerns in a recommender system. Such a mechanism has rarely been documented in prior work. Specifically, we focus on two constructs—*value of disclosure* and *trust*—as potential explanations for privacy concerns about different types of information in a recommender context.

Through exploratory factor analyses, the current study divided 21 pieces of demographic information into identifiable vs. unidentifiable information, and divided 26 types of products into sensitive vs. non-sensitive categories. Based on the preliminary categorization, we tested the effects of user control and data input in a popular recommender system (i.e., Amazon.com) via an online experiment with four different scenarios. We found a significant influence of control on reducing users' concerns about both types of information. For product-related information, if data is accessed by the recommender implicitly (i.e., through purchase history), the presence of user control plays a significant role in decreasing privacy concerns. However, if product data is collected by the recommender explicitly (i.e., through product rating), user control does not help to alleviate users' privacy concerns over information releasing. In addition, we found that value of disclosure, rather than trust, explains the underlying psychological mechanism of control's influence on privacy concerns.

This paper makes three main contributions to privacy research in recommender systems. First, we created two item-based information indices based on users' privacy concerns (i.e., one for demographics and one for product-related information). Our indices extended previous research [27, 47] by extracting new factors. Based on these categorizations, future system developers can strategically adopt data input methods and privacy protection solutions. Second, our findings showed that user control was effective in reducing privacy concerns for implicit data (i.e., purchase history), but not for explicit data (i.e., product rating), which casts doubt on the current trend of embedding control mechanisms unconditionally for privacy-concern reduction. Thus, the implementation of a control mechanism in recommender systems should also be designed strategically. Last, adding to existing research on privacy in recommender systems, we propose *perceived value of information disclosure* as a psychological mechanism that could explain the phenomenon. We then discuss practical implications and directions for future research.

2. THEORETICAL BACKGROUND AND HYPOTHESES

Substantial research has highlighted the issue of privacy concern in recommender systems [3, 11]. In this section, we first review prior work on personalization and the relationship between data input type and privacy concerns. We then discuss the effectiveness of affording user control in alleviating such concerns in recommenders. Last, we consider psychological mechanisms that might explain users' privacy concerns. Built on extant previous privacy research, we also propose our hypotheses, research questions, and conceptual model.

2.1 Personalization and Privacy Concerns

Personalization, or proactive tailoring of products and services based on individuals' preferences and needs [9], is at the heart of recommender systems' functionality and technology [8]. It is of great importance to online vendors because user information can help them predict demand, build customer loyalty, and increase cross-selling possibilities [39]. Personalization also has been found to be of significant value to users, by providing convenience and better service matching [9], saving time and effort, and promoting an optimal user experience [28]. However, users may be hesitant to savor the benefits brought by sophisticated personalization technology [13] because these benefits inherently come with the sacrifice of some privacy. For example, personalized convenience may rely on unsolicited data collection [33], or the fact that recommender systems share user data with third parties [6]. This phenomenon is known as the "privacy-personalization tradeoff" [3, 9]. Some studies suggest that users rationally calculate the net value gained from information disclosure accounting for privacy loss [15, 52], whereas others argue they superficially process personalization cues on an interface [45, 54].

Regardless of how personalization is interpreted by users, its effectiveness mainly depends on two factors: the recommender's ability to capture and analyze user data, and users' willingness to share data and use personalized services [9]. The former may refer to different ways of collecting data (i.e., explicit vs. implicit data input) and also different types of data (i.e., demographic vs. product-related information); the latter points to an individual characteristic, namely the extent to which one values information disclosure in return for personalized benefits [25, 52]. Both aspects are addressed in this study.

2.2 Data Input in Recommender Systems

The efficacy of a successful recommender system is achieved by extensively acquiring, storing, and processing user data. This data varies in sensitivity and is gathered through different approaches. This study investigates users' privacy concerns regarding individual information items (i.e., demographic vs. product) used by a recommender system, and the influence of data input (i.e., explicit vs. implicit) on users' privacy concerns.

2.2.1 Demographic vs. Product Information

Recommender systems collect and analyze static demographic information that can be linked to individual identities (e.g., email addresses, social security numbers) [27] and dynamic online behavioral data that can infer one's tastes and preferences (e.g., purchase history, product ratings) [47]. Although this information has been deemed significant in affecting user privacy [19, 40], variation among individual information items in triggering privacy concerns has not been explored. Considering the vast number of footprints users leave online every day, and the difficulty in balancing prediction accuracy and privacy protection, it is critical to identify different types of user data that vary in sensitivity so that system developers can strategically implement different protection mechanisms.

There are two broad types of online user data: static demographic data and dynamic behavioral data. Past research has labeled them as "demographics" and "context" [27], corresponding to two recommendation strategies: content filtering and collaborative filtering [30]. Behavioral data (i.e., recording what a user browses, clicks, and purchases online) is always associated with specific products. In an online shopping scenario, 23 product items were identified to raise different levels of privacy concerns, leading to reluctance in purchasing them [47]. Based on these previous definitions and findings, this study conducts an item-based privacy concern rating and analysis to evaluate how information type is connected with privacy concerns. Hence, we propose the following research questions:

RQ1a: What types of demographic information used by recommender systems for personalized recommendation will trigger privacy concerns?

RQ1b: What types of product information used by recommender systems for personalized recommendation will trigger privacy concerns?

2.2.2 Explicit vs. Implicit Data Input

To provide personalized recommendations, recommender systems rely on two kinds of user data input: explicit and implicit [30]. Past research has labeled them in various ways, for example, pull vs. push [48], overt vs. covert [52], customization vs. personalization (i.e., agentic actions vs. tailoring) [44], to name a few. Explicit input is direct feedback from users that clearly expresses their preferences and tastes, such as product ratings and movie critiques [10, 19]. Implicit input, on the other hand, is information unconsciously left by users online, which is often clustered automatically by algorithms to identify user-item connections for future recommendations [20]. Implicit input includes browsing history, purchase history, clicking behaviors, and search patterns [30]. Although users do not explicitly express their opinions for implicit data input, their tendencies can often be speculated based on their behavioral patterns. Explicit input requires users to be willing to give out information consciously, whereas user effort is not necessary for implicit input [40]. Therefore, the main difference between these two approaches lies in the presence of user consciousness and initiative. The two types of approaches are often adopted simultaneously in recommender systems for better prediction accuracy and efficiency [30].

Previous research has examined the effectiveness and impact of these input types on user perceptions and behaviors from different perspectives. For example, explicit input, rather than implicit input, has been found to be preferable in location-based advertising because users perceived more control and benefits in it and would also be more likely to employ it [48]. Implicit input may appeal to online vendors because they do not need to lobby users to opt in and because it may stimulate impulsive purchasing [48]. To users, however, implicit input can be intrusive because it means their data is tracked without consent; this could diminish the perceived value of recommendations, and even trigger negative reactions such as avoidance [16] and privacy concerns [7, 48]. In this study, we consider product rating and purchase history as representations of explicit and implicit data input, respectively. Drawing on this previous work about users' negative perceptions of implicit input, we posit the following hypothesis:

H1: In a recommender system, implicit data input (i.e., purchase history) will trigger greater privacy concerns about product information than explicit data input (i.e., product rating).

2.3 Empowering Users with Control to Reduce Privacy Concerns

Researchers, system developers, and policy makers have been creating solutions at all levels to cope with rising privacy concerns in online recommender systems and minimize the compromise of prediction accuracy. For example, Heitmann et al. [22] proposed an architecture that enables users to decide what personal information can be accessed by which service providers; Arlein et al. [2] designed a data protection mechanism that allows users to hide their real identities and use personae for information sharing; Xu et al. [53] demonstrated that privacy assurance approaches such as the TRUSTe seal can also reduce users' privacy concerns by way of perceived control over personal information. Most of these solutions endow users with either the capability of actually controlling their information sharing or with a perceived sense of control, which has been found effective in alleviating privacy concerns.

Indeed, the idea of privacy is often associated with control over personal information; if something is considered to be private, we want to be able to protect it [12, 43]. Many researchers directly define privacy as a sense of control [35, 49]. The notion of control is frequently studied as a key factor of privacy concern [34]. Loss of control over collection and usage of information has been found to lead to a greater sense of privacy invasion among online consumers [14, 41]. Milne and Boza [37] showed that, in general, individuals have less privacy concerns when they have a greater sense that they controlled the disclosure and subsequent use of their personal information. Acquisti and Gross [1] found that Facebook users who were not concerned about privacy of the information they posted online also felt a greater sense of control over it. Given the negative relationship between control and privacy concerns suggested by prior research, we propose the following hypotheses:

H2a: The presence of user control will lead to a decreased level of privacy concern about demographic information compared to the no-control condition.

H2b: The presence of user control will lead to a decreased level of privacy concern about product information compared to the no-control condition.

Because we also consider the effect of data-input type (i.e., explicit and implicit), which only applies to product-related information, we further ask the following research question:

RQ2: Is there an interaction effect between data input type and user control on privacy concerns toward product information?

2.4 Psychological Mechanisms of Privacy Concerns

Although the work discussed above provides insightful findings, it remains unclear which particular psychological mechanisms determine privacy concerns in a recommender context. Given numerous technological attempts in affording user control to reduce privacy concerns, we employ *perceived value of information disclosure* and *trust toward the recommender system* as potential underlying psychological paths.

2.4.1 Value of Disclosure to Users

The privacy calculus model posits that perceived value in information disclosure is often evaluated by weighing benefits and risks [14]. In a privacy context in recommender systems, then, perceived value of information disclosure can be defined as the trade-off between what users can gain from using the recommender and what risks users need to take in disclosing their information [52]. The deployment of a user control mechanism in a recommender system is likely to increase users' perceived benefits, as well as reduce their concerns about privacy loss. This leads to the following hypothesis:

H3: The presence of user control will lead to greater perceived value of information disclosure compared to the lack of control.

We define perceived value of information disclosure as the trade-off between benefits received from the recommender system and privacy loss to the system; therefore, the greater the perceived value, the more perceived benefits outweigh privacy risks, and the less likely one is to be concerned about privacy. As such, we further posit the following hypothesis:

H4: Perceived value of information disclosure mediates the relationship between user control and privacy concerns toward specific information.

2.4.2 A Trust-building Mechanism

Trust also has been found to be a concept that is closely associated with privacy in an online environment [18, 42]. Belanger et al. [5] found that consumers heavily rely on the trustworthiness of an online vendor to disclose their information; Milne and Boza [37] demonstrated that trust-building is more effective than concern-reducing in managing online users' information. Studies also showed that, at an institutional level, trust can significantly mediate the effect of privacy assurance practices on users' privacy concerns [14, 51]. Based on prior research, privacy coping strategies (e.g., providing user control) may assure users that their information will only be accessed and used with their consent, thereby inducing perceived trust toward the service provider, and eventually leading to a reduced level of privacy concern. Thus, we hypothesize the following:

H5: The presence of user control will lead to greater perceived trust toward the recommender system compared to the lack of control.

H6: Perceived trust toward the recommender system mediates the relationship between user control and privacy concerns.

2.4.3 Influences of Personal Traits

Personal traits, or individual characteristics, reflect human natures and can determine one's perceptions and behavioral patterns in many situations [38]. This is especially true in a privacy context because individual dispositions are often linked with one's privacy concerns and tendency to disclose information [4]. Three personal traits are particularly of relevance to this study: general privacy concern, perceived value of personalization, and perceived importance of control. To account for their potential influences on users' privacy concerns, these traits are all included as control variables.

2.5 A Conceptual Model

Grounded in theoretical research and prior empirical studies, we propose a conceptual model for the current study (Figure 1).

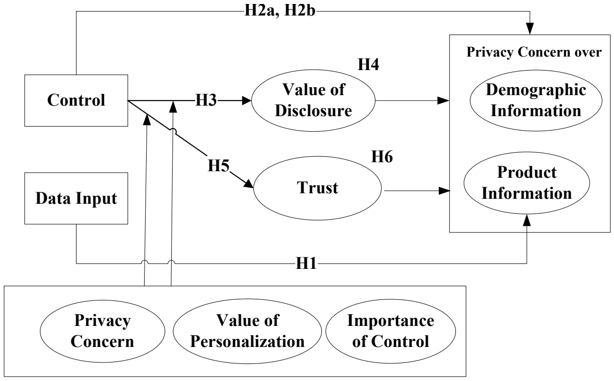

Figure 1. A conceptual model

3. METHODOLOGY

3.1 Study Design

The study's design consists of two components: item-based privacy concern ratings and scenario-based privacy concern probing. On the whole, a 2 (data input: explicit vs. implicit) x 2 (user control: presence vs. absence) scenario-based online user study was conducted to answer our research questions and test our hypotheses and conceptual model. To provide an index of information items that vary in sensitivity and differentiate them by degree of user privacy concern, we included 21 pieces of demographic information inspired by Knijnenburg et al. [27] and 26 product types [47]. Based on the scenario, participants were asked to rate how concerned they would be if Amazon.com accessed different types information for personalized recommendations. The purpose of this rating was to provide a relative measure of privacy concern on an item-by-item basis, rather than an absolute scale of information sensitivity.

3.2 Participants and Recruitment

We recruited all participants (N = 385) through Amazon Mechanical Turk (MTurk, www.mturk.com), a recruitment source that has become popular for conducting online user studies in recent years [26]. We restricted participants to US residents with a North American IP address and a Human Intelligence Task (HIT) approval rate of 90% or higher. Participants were also required to have made at least one purchase on Amazon.com in the past year so that the scenario setting would seem applicable to them. As an incentive, we paid each eligible participant 50 cents for a completed task. The majority of the participants were male (63.2 %) and Caucasian (75.8 %). The average age was 31.14 (SD = 10.70). We recognize the potential confounding effect of using an Amazon-based participant pool for an Amazon-related study. However, any confounding effects will be identical across conditions, so this should not cause any analytic problems.

3.3 Scenarios

We created four scenarios to examine the effects of user control (presence vs. absence) and data input (explicit vs. implicit) on users' privacy concerns in a recommender system. All four scenarios were grounded in an online shopping context with Amazon.com due to its prominent role in recommender systems. Specifically, we instructed participants to imagine that they had purchased all of the listed products from Amazon.com. In the scenarios with presence of user control, participants were explicitly told that they had control over the extent to which Amazon.com could access their demographic information and purchase history in exchange for personalized recommendations; in the scenarios without user control, such information was not provided. Within these scenarios, we also varied two types of product-related data input by asking participants to evaluate their level of concern over releasing product information in their purchase history (i.e., implicit data input) or product ratings (i.e., explicit data input) to Amazon.com for personalized recommendations. Participants were randomly assigned to one of the four conditions/scenarios.

3.4 Procedure

After participants were pre-screened for eligibility, they were randomly assigned to one of the two user control scenarios (i.e., presence vs. absence), where they were instructed to evaluate their levels of concern over releasing 21 types of demographic information to Amazon.com in exchange for personalized recommendations. After that, participants were randomly assigned to one of the two data input scenarios (i.e., purchase history vs. rating), where they were asked to assess their privacy concerns over releasing 26 types of products purchased from Amazon.com in exchange for personalized recommendations. After the item-based privacy concern evaluations, we measured perceptual variables regarding users' attitudes toward the recommender system and individual characteristics. We then collected demographic information for use as control variables.

3.5 Measurements

To the extent possible, we adopted measurement scales for the main constructs in this study from prior research to fit the Amazon.com recommender context.

Inspired by Knijnenburg et al. [27], we included 21 pieces of demographic information that vary in sensitivity (e.g., email addresses, phone numbers, social security numbers). For product items, we included 26 products that also differ in sensitivity (e.g., textbook, hunting knife, bulletproof jacket) [47]. Privacy concerns about these items were measured on a Likert-type scale from "1 = not concerned at all" to "7 = extremely concerned."

Value of information disclosure was assessed by 3 items adapted from Kim et al. [25] and Xu et al. [52] (e.g., "The value I gain from use of Amazon.com's service is worth the information I give away.") (α = .756). Trust toward Amazon.com was measured with 6 items (e.g., trustworthy) (α = .886) [36].

In terms of individual differences, we measured participants' general privacy concern, perceived value of online personalization, and perceived importance of control. General privacy concern was measured via 3 items (e.g., "I am sensitive about giving out information regarding my preferences") [9] (α = .828). Value of online personalization was assessed with 6 items (e.g., "I value websites that are personalized for my usage experience preferences.") derived from Chellappa and Sin [9] (α = .855). Participants were also asked to indicate their perceived importance of control in the recommender context (e.g., "It is important for me to control the amount of information accessed by Amazon.com for personalized recommendations") (α = .943). This final measure was specific to the study and, therefore, created by the researchers.

All these measures took the form of 7-point Likert-type scales, with 1 being the lowest level and 7 the highest. A complete list of measurement items can be found in Appendix C.

4. RESULTS

We present our results by first describing the item-based analyses of privacy concerns over demographic and product information in response to our research questions. We then examine effects of data input and user control on participants' psychological perceptions and privacy concerns. Finally, we test our conceptual model of the psychological mechanism of privacy concerns in a recommender system via mediation analysis and structural equation modeling

To rule out confounding issues, we statistically control for general privacy concern, perceived value of online personalization, and perceived importance of control, along with other demographic information (e.g., gender, age, education).

4.1 Item-based Analyses of Privacy Concern

To discriminate sensitive information items from non-sensitive items and create an index of data types based on users' privacy concerns, we performed exploratory factor analyses (EFA). Table 1 shows a complete list of the 21 pieces of demographic information we included in the study, and Table 2 shows the 26 specific product types.

4.1.1 Demographic Information Type: Unidentifiable vs. Identifiable

The 21 items regarding privacy concerns over demographic information were first subjected to a principal axis factoring analysis (PAF) with an oblique, promax rotation. PAF was chosen because it generally produces outcomes close to maximum likelihood extraction and it is not overly sensitive to nonnormality [17]. An examination of the Kaiser-Meyer Olkin (KMO) measure of sampling adequacy suggested that the sample was factorable (KMO = .951). Scree-plot analysis indicated two factors for demographic information. The rotated pattern matrix of the item pool is shown in Table 1. One severely cross-loading item, *date of birth*, was dropped from the analysis based on the 0.3 rule (i.e., an item's highest loading should be at least 0.3 higher than its other loadings).

The 12 types of personal information that loaded onto Factor 1 represent general personal attributes that cannot be used as identifiers of a particular person. Hence, we labeled Factor 1 as "unidentifiable demographic information." On the contrary, the 8 items that loaded onto Factor 2 represent unique information that can be used to identify or locate an individual. Therefore, this was labeled as "identifiable demographic information." The individual items for each factor, factor loadings, and reliabilities can be found in Table 1.

To address RQ1a and test the difference in causing privacy concerns between the two demographic information types, a paired samples *t*-test was conducted. Results showed that users were significantly more concerned about releasing identifiable demographic information (M = 3.850, SD = 1.571) to Amazon.com in exchange for personalized recommendations than unidentifiable demographic information (M = 3.188, SD = 1.469, p < .001).

Table 1. Exploratory factor analysis and privacy concern levels for demographic information

Component	Privacy Concern	Factor Loading	
	(Range: 1 to 7)	1	2
Unidentifiable Demographic Information (α = .938)			
Education	2.95	**.960**	-.170
Relationship	3.22	**.913**	-.177
Race	2.68	**.859**	-.097
Field of work	3.06	**.846**	-.005
Tech use	3.15	**.742**	.040
Interest	2.70	**.731**	.031
Gender	2.40	**.710**	.089
Age	2.65	**.688**	.174
Company	3.25	**.638**	.173
Calendar	3.66	**.604**	.040
Income	3.93	**.552**	.242
Web browsing	4.60	**.486**	.125
Identifiable Demographic Information (α = .904)			
Credit card	4.73	-.270	**.920**
Home address	3.57	-.088	**.919**
Phone number	3.88	.017	**.830**
Email	3.05	.051	**.756**
Name	2.94	.169	**.631**
Location	3.44	.232	**.616**
IP address	3.86	.113	**.585**
SSN	5.79	.022	**.463**
Dropped Item			
Date of birth	3.38	.384	.466

4.1.2 *Product Type: Non-sensitive vs. Sensitive*

In a similar manner, the 26 specific products tested were subjected to a PAF with a promax rotation. KMO suggested that the sampling was adequate for factor analysis (KMO = .967). Scree-plot analysis indicated two distinct factors for product types. The rotated pattern matrix is in Table 2. *Cigarette*, *lingerie*, and *bulletproof jacket* were dropped because of cross-loading, and *shoes* was discarded because of multicollinearity.

The 12 types of products that loaded onto Factor 1 are all products that are normally not considered to be sensitive, such as *office supplies* and *everyday necessities*. This factor was labeled as "non-sensitive products." The 13 types of products that loaded onto the second factor are products related to personal values and mental states, such as *HIV tests*, *depression-related books*, *bomb-making books*. Thus, we labeled Factor 2 as "sensitive products." Individual items for each factor, factor loadings and reliabilities can be found in Table 2.

To address RQ1b and examine how sensitive products are different from non-sensitive products in triggering privacy concerns, we conducted a paired samples t-test. Results showed that users' were significantly more concerned about releasing information about sensitive products (M = 3.762, SD = 1.768) to Amazon.com in exchange for personalized recommendations than non-sensitive products (M = 2.085, SD = 1.350, p < .001).

Table 2. Exploratory factor analysis and privacy concern levels for product information

Component	Privacy Concern	Factor Loading	
	(Range: 1 to 7)	1	2
Non-sensitive Products (α = .965)			
Furniture	1.83	**.972**	-.126
Food	1.85	**.969**	-.108
Flower	1.83	**.959**	-.099
Laptop	2.00	**.939**	-.064
Textbook	1.87	**.926**	-.055
Game	1.92	**.901**	-.046
Jewelry	2.07	**.867**	-.003
Peroxide	2.15	**.767**	.108
Hunting knife	2.35	**.678**	.229
Fertilizer	2.26	**.657**	.143
Weight loss product	2.48	**.614**	.303
Sensitive Products (α = .956)			
STD medication	4.61	-.248	**.978**
HIV test	4.58	-.173	**.942**
Sex toy	4.26	-.130	**.920**
Porn DVD	4.29	-.107	**.902**
Adult diaper	3.80	-.027	**.822**
Lubricant	3.37	.126	**.769**
Book-Bomb making	4.59	-.114	**.756**
Pregnancy Test	3.48	.139	**.732**
Book-Depression	3.23	.164	**.714**
Condom	3.26	.226	**.679**
Book-Bankruptcy	3.28	.220	**.655**
Dropped Items			
Shoes	1.83	1.011	-.175
Cigarette	2.66	.501	.349
Lingerie	2.95	.366	.526
Bulletproof Jacket	3.22	.309	.511

4.2 Effects of Data Input and User Control

Based on the level of privacy concern, users' demographic information can be classified into two categories: unidentifiable demographic information and identifiable demographic information. Similarly, product types can be classified into two categories: non-sensitive products and sensitive products. We adopt these classification results in the following analyses.

4.2.1 *Effects of Data Input*

In order to test the effects of user data input (explicit/rating vs. implicit/purchase history) in the recommender system on users' perceived privacy concerns, a series of analyses of covariance (ANCOVAs) were conducted, controlling for the influences of general privacy concern, perceived value of personalization, perceived importance of control, and demographics (e.g., age, gender, education). It is worth noting that data input type in recommender systems only applies to product-related information, not demographic information; product information can be obtained through user rating or history checking, whereas demographic information can only be obtained through user input. Results showed a significant main effect of data input on value of information disclosure about product-related items, F (1, 376) = 7.85, p < .01. Specifically, participants perceived more value in disclosing purchasing history (M = 4.528, SD = .078) than in disclosing product ratings (M = 4.234, SD = .077) in exchange for personalized recommendations. However, data input's effect on privacy concerns about product information was not significant. Thus, H1 was not supported.

4.2.2 Effects of User Control

A multivariate analysis of covariance (MANCOVA) was conducted to investigate the effects of user control in the recommender system. Results indicated a significant overall main effect of user control, Wilks' Λ =.941, *F* (1, 372) = 5.839, *p* < .001. Subsequent univariate analyses showed that participants tended to express a higher level of perceived value of information disclosure [(*M* = 4.519, *SE* = .077), *F* (1, 376) = 6.629, *p* = .01] and significantly less concern about their unidentifiable demographic information [(*M* = 2.934, *SE* = 1.317), *F* (1, 376) = 7.863, *p* = .005], identifiable information [(*M* = 2.934, *SE* = 1.417), *F* (1, 376) = 22.345, *p* < .001], non-sensitive products [(*M* = 1.914, *SE* = 1.173), *F* (1, 376) = 4.752, *p* = .030], and sensitive products [(*M* = 3.492, *SE* = 1.714), *F* (1, 376) = 6. 352, *p* = .013], when they had control over information access then not (*M* = 4.236, *SE* = .078; *M* = 3.369, *SE* = .090; *M* = 4.178, *SE* = .097; *M* = 2.230, *SE* = .093; *M* = 3.974, *SE* = .118, respectively) (see Figure 2 (a) & (b)). However, the effect of control on perceived trust toward the recommender was not significant (*F* (1, 376) = .038, *p* = .846). Therefore, H2a, H2b, and H3 were all supported, but H5 was not.

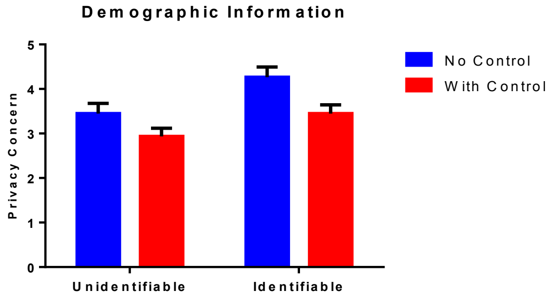

Figure 2(a). Effects of user control on privacy concerns about demographic information

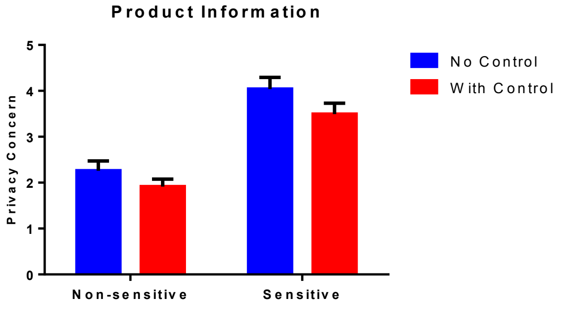

Figure 2(b). Effects of user control on privacy concerns about different information types

4.2.3 Interaction Effects between Data Input and User Control

To answer RQ2, we tested the interaction effects between data input and user control. We found that data input type significantly moderated the relationship between the existence of user control and perceived privacy concern about non-sensitive product information, *F* (1, 374) = 4.657, *p* = .032, but not sensitive product information, *F* (1, 374) = 1.691, *p* = .154. Specifically, empowering users with control over information release significantly lowered their concerns over purchase history containing non-sensitive products (*M* = 2.402, *M* = 1.914, for no-control and control conditions respectively). However, if users were asked to explicitly rate the non-sensitive products they had purchased before, such a control would not make a difference (*M* = 2.125, *M* = 2.133, for no-control and control conditions respectively) (Figure 3). This significant interaction effect did not exist for sensitive products.

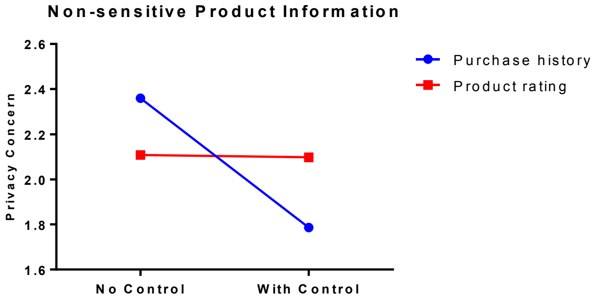

Figure 3. Interaction effect of user control and data input on privacy concerns about non-sensitive products

4.3 Testing the Conceptual Model

Before testing the overall conceptual model, we first speculated the degree to which effects of user control on privacy concerns over the four types of information (i.e., unidentifiable demographics, identifiable demographics, non-sensitive products, sensitive products) might be mediated by the two proposed psychological mechanisms—perceived value of disclosure and trust. An SPSS script developed by Hayes [21] was adopted to probe such mediation effects. As shown in Figure 4, perceived value of disclosure significantly mediated the effects of control on privacy concerns about unidentifiable demographic information (*β* = -.14, *p* < .001, Figure 4a), identifiable demographic information (*β* = -.12, *p* < .001, Figure 4b), non-sensitive products (*β* = -.08, *p* < .001, Figure 4c), and sensitive products (*β* = -.16, *p* < .001, Figure 4d). However, perceived trust toward the recommender system was not a significant mediator in any of these relationships. All path coefficients are shown in Figure 4. These findings provide support for H4 but not for H6.

Because we did not find any significant effect of user control on trust toward the recommender, nor did we yield a significant indirect effect of user control on privacy concerns via trust, we removed the trust construct from our conceptual model for the following statistical testing.

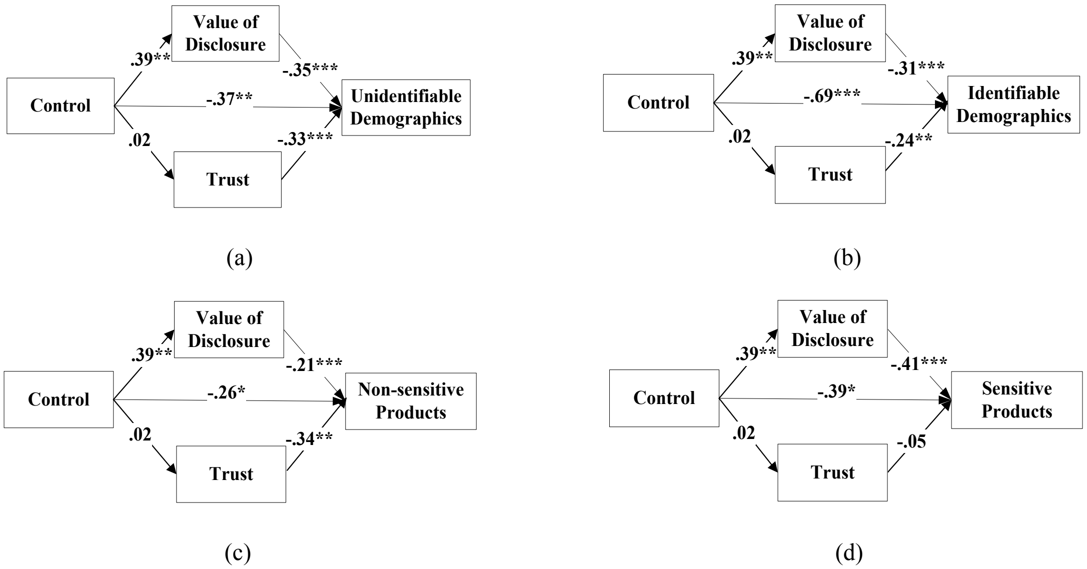

Figure 4. Path models for control's effect on privacy concerns about information [four types, (a)-(d)] with value of disclosure and trust as possible mediators (**p*< .05, ***p*< .01, ****p*< .001**)**

Given that user control plays an important role in affecting privacy concerns through perceived value of information disclosure, we tested the overall conceptual model with structural equation modeling (SEM) to map out the relationships among our main constructs. The 8-latent-factor structure with 57 individual items was found to retain a reasonably good fit: χ^2 = 9293.596, *df* = 3884, *p* < .001, root mean square error of approximation (RMSEA) = .049, 90% confidence intervals (CI): .048-.050, comparative fit index (CFI) = .825. And a subsequent multigroup structural equation modeling (MGSEM) with data input type (explicit vs. implicit) as the grouping variable yielded close good-fitting models. Figure 5 presents the final overall model and standardized path coefficients.

Consistent with previous findings, empowering users with control in the recommender system tends to enhance participants' perceived value of information disclosure. Such increased value of disclosure directly alleviates users' concerns over releasing their demographic and product-related information in exchange for personalized recommendations.

To further probe this effect, bootstrapping procedures were employed using 2000 bootstrap samples and a bias-corrected confidence interval in a multigroup analysis. With data input type as the grouping variable, results showed that the significant mediating effects of perceived value of disclosure only exist when data input is implicit (i.e., when purchase history is accessed for personalized recommendations). Specifically, perceived value of disclosure significantly mediated the relationship between the presence of user control in the recommender and privacy concerns about non-sensitive products (β = -.13, *p* = .006) and sensitive products (β = -.12, *p* = .009) when users thought their purchase history would be accessed. However, in the product-rating scenario, such mediating effects were not significant (β = -.06, *p* = .16; β = -.06, *p* = .12; for non-sensitive products and sensitive products respectively).

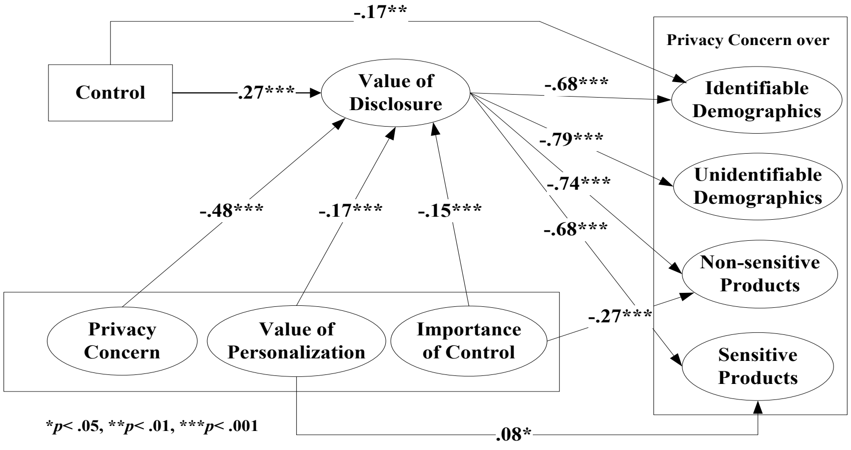

Figure 5. SEM explaining the psychological mechanism of privacy concerns

In sum, our findings suggest that different types of user information in a recommender system should be treated differently depending on the degree of privacy concern they may trigger. A user control mechanism is effective in reducing the concern regarding implicit data input only. In addition, control influences privacy concerns about both demographic and product-related information by way of users' perceived value of information disclosure.

5. DISCUSSION AND CONCLUSIONS

In this section, we provide interpretations of the study's main findings, present design suggestions for recommender systems, and then discuss the limitations and directions for future work.

5.1 Interpretation of Results and Design Implications

5.1.1 User Data Categorization and Sensitivity Ranking

The findings of our study suggest that users' online information is multi-dimensional regarding privacy concerns, especially in a recommender context. Although this seems self-explanatory, it is often neglected in privacy research and recommender system design. Specifically, demographic information that is frequently required for online service registration can be divided into two categories: unidentifiable information and identifiable information. Unidentifiable information consists of items describing one's personal attributes (e.g., age, gender) that cannot be used to uniquely pinpoint the individual, whereas identifiable information is more accurate in pointing to the individual's identity exclusively (e.g., phone numbers, email addresses). People are significantly more concerned about the recommender system accessing their identifiable information than their unidentifiable information. In a similar manner, product items can be broadly grouped into non-sensitive types and sensitive types. Users are significantly more worried about their previous purchases of sensitive products (e.g., adult diapers, HIV tests) being accessed for personalized recommendations than they are about their previous purchases of non-sensitive products (e.g., jewelry and shoes).

These item-based analyses and categorizations provide a relative information-ranking system in terms of privacy concern in recommender systems, thus refining existing research on general privacy concern about user information. Although a few previous studies have also identified specific information items that vary in sensitivity in recommender systems [27, 47], the current categorization extended prior research by extracting new factors, which can be used as a reference in future studies and designs. These new factors suggest that recommender system designers should treat users' information discriminatively and strategically based on their levels of sensitivity for pattern prediction and personalized recommendations. Algorithm developers should be well aware of what information users are more hesitant to disclose, so as to adjust the degree of information tracking and use, as well as to provide appropriate coping strategies. In line with the "privacy-personalization trade-off," unsolicited access to users' sensitive information may trigger severe privacy concerns that could affect users' overall experiences [28]; therefore, identifiable and sensitive data should be more cautiously handled in exchange for prediction accuracy. As a design suggestion, recommender systems should introduce user control or privacy assurance mechanisms to help alleviate users' privacy concerns. Also, user data with different sensitivity levels (e.g., identifiable vs. unidentifiable information) can be potentially protected with different levels of privacy remedies.

5.1.2 Effectiveness of User Control and Data Input Type

We also showed that the presence of a user control mechanism over information disclosure greatly impacts users' privacy concerns in a recommender system, which is consistent with previous findings [23, 24]. For demographic information, user control significantly lowered privacy concerns. However, for product-related information, such effects pertain to non-sensitive products only and are significantly moderated by data input type (i.e., explicit vs. implicit). When personalized recommendations are provided based on one's purchase history (i.e., implicit input), users tend to feel concerned about what they have bought when they have no control, but they feel significantly more relieved if they have control over information access by the service provider. This may be due to a sense of intrusiveness; implicit data input is often unsolicited, so users do not always expect that such information will be used for recommendation purposes. Affording users control over information release would not only allow users to modify their privacy settings and gain a sense of autonomy, but also help them predict what information might be at risk, thereby reducing the concern level resulting from uncertainty.

However, if users are explicitly told to rate the products they have purchased before (i.e., explicit input), the control mechanism does not help much in alleviating their concerns (Figure 3). Even though the control mechanism allows users to manage what information could be accessed by the recommender system, it seems that the control mechanism works for implicit data input rather than explicit data input. As discussed, implicit data input can trigger a sense of intrusiveness because records are often traced without users' permissions. On the other hand, explicit data input (i.e., product rating) is initiated by users, themselves, so users are already imbued with a sense of competency; because of this, an extra control mechanism would probably not change their perceptions. If users felt concerned about expressing their opinions and exposing their preferences, they would be unlikely to rate the products in the first place.

Furthermore, this intriguing interaction effect exists for non-sensitive products, but not for sensitive products. It could be that users are generally confident in protecting information related to non-sensitive products they have purchased, and the addition of a control mechanism further strengthens this confidence. However, when it involves sensitive products, users become much more cautious that their concern level may reach a "ceiling effect." Therefore, neither data input type nor the presence of control can alleviate the heightened concern levels.

This is the most intriguing finding of the current study, which casts doubts on ongoing efforts to embed user control in all recommender systems. The current study suggests that, for operations that do not require users' conscious attention and actions (e.g., tracking and analyzing their purchase history), an active control mechanism is needed to overcome perceived intrusiveness and privacy concerns. However, users are already empowered with deliberation in an explicit rating situation, thus the extra control could seem redundant. Also, a control mechanism may only be convincing enough to protect information about non-sensitive products. As a design implication, a user control mechanism may not be as effective for recommendations relying on explicit data input, compared to those based on implicit user data input. Additionally, users seem to have persistent concerns about previously purchased sensitive products, and this cannot be easily mitigated by control mechanisms. There also is an asymmetric information problem between the service provider and the user—a lack of awareness could be another cause of the current finding. That is, users might not be aware of what companies can do with their non-sensitive information. Increasing the awareness level may boost the effectiveness of user control mechanisms. Therefore, system designers should carefully weigh the advantages and disadvantages of a control mechanism in addressing privacy concerns depending on the data input type, data sensitivity levels, and the existence of an awareness mechanism.

5.1.3 Psychological Mechanism of Privacy Concern

This study also proposed and tested a conceptual model for demonstrating the underlying psychological mechanisms of privacy concerns in a recommender system. Our findings showed that, after controlling for individual differences, users' perceived value of information disclosure explains how user control affects privacy concerns. In the current study, perceived value of information disclosure is measured based on the privacy calculus model, representing a trade-off between perceived benefits gained from personalized recommendations and risks of privacy invasion. Our results suggest that the mere mention of a control mechanism in the recommender system scenarios can elevate perceived value of information disclosure. This is likely because the addition of control boosts users' perceived value of the entire system, so users are more confident about trading in their privacy for personalized services. This heightened perceived value leads to lesser privacy concerns. Because perceived value often comes from perceived usefulness and effectiveness of the system [25], recommender system designers should focus on these aspects to improve users' psychological evaluation of the system so as to conquer privacy concerns. This is yet another motivation for designers to strive for a better recommender system with efficient functionality.

5.2 Limitations and Future Research

Although the current scenario-based design has its merits in many aspects, especially in an exploratory study, our manipulation and setting of the main constructs (e.g., control, data input) relied solely on users' assumptions and imaginations as instructed by our study descriptions. Participants may have had a different impression and evaluation of a recommender system if they could interact with a real interface. Their concerns over various information types also depended on a hypothetical picture of what they had previously purchased from Amazon.com. Therefore, the scenario-based design may lack external validity. Future research could implement a real interface prototype based on our preliminary findings, examine users' real behaviors (e.g., purchasing, rating) in a natural setting over a longitudinal period, and then measure their privacy concern levels. In addition, apart from perceived value of disclosure, other psychological mechanisms of privacy concerns in recommender systems should also be explored.

6. REFERENCES

[1]Acquisti, A. and Gross, R. 2006. Imagined communities: Awareness, information sharing, and privacy on the Facebook. In Proceedings of the Privacy enhancing technologies. 36-58.

[2]Arlein, R.M., Jai, B., Jakobsson, M., Monrose, F., and Reiter, M.K. 2000. Privacy-preserving global customization. In Proceedings of the 2nd ACM conference on Electronic commerce. 176-184.

[3]Awad, N.F. and Krishnan, M. 2006. The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to be Profiled Online for Personalization. MIS quarterly, **30**(1).

[4]Bansal, G., Zahedi, F., and Gefen, D. 2010. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decision Support Systems, **49**(2): 138-150.

[5]Belanger, F., Hiller, J.S., and Smith, W.J. 2002. Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. The Journal of Strategic Information Systems, **11**(3): 245-270.

[6]Bennett, J. and Lanning, S. 2007. The netflix prize. In Proceedings of the KDD cup and workshop. 35.

[7]Berendt, B. and Teltzrow, M. 2005. Addressing users' privacy concerns for improving personalization quality: Towards an integration of user studies and algorithm evaluation, in Intelligent Techniques for Web Personalization. Springer. 69-88.

[8]Burke, R. 2002. Hybrid recommender systems: Survey and experiments. User Modeling and User-Adapted Interaction, **12**(4): 331-370.

[9]Chellappa, R.K. and Sin, R.G. 2005. Personalization versus privacy: An empirical examination of the online consumer's dilemma. Information Technology and Management, **6**(2-3): 181-202.

[10]Chen, L. and Pu, P. 2012. Critiquing-based recommenders: survey and emerging trends. User Modeling and User-Adapted Interaction, **22**(1-2): 125-150.

[11]Cranor, L.F. 2004. I didn't buy it for myself, in Designing personalized user experiences in eCommerce. Springer. 57-73.

[12]Culnan, M.J. 1993. " How Did They Get My Name?": An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use. MIS quarterly, **17**(3).

[13]Culnan, M.J. 2000. Protecting privacy online: Is self-regulation working? Journal of Public Policy & Marketing, **19**(1): 20-26.

[14]Culnan, M.J. and Armstrong, P.K. 1999. Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. Organization science, **10**(1): 104-115.

[15]Culnan, M.J. and Bies, R.J. 2003. Consumer privacy: Balancing economic and justice considerations. Journal of social issues, **59**(2): 323-342.

[16]Edwards, S.M., Li, H., and Lee, J.-H. 2002. Forced exposure and psychological reactance: Antecedents and

consequences of the perceived intrusiveness of pop-up ads. Journal of Advertising, **31**(3): 83-95.

[17]Finch, J.F. and West, S.G. 1997. The investigation of personality structure: Statistical models. Journal of Research in Personality, **31**(4): 439-485.

[18]Friedman, B., Khan Jr, P.H., and Howe, D.C. 2000. Trust online. Communications of the ACM, **43**(12): 34-40.

[19]Gena, C., Brogi, R., Cena, F., and Vernero, F. 2011. The impact of rating scales on user’s rating behavior, in User Modeling, Adaption and Personalization. Springer. 123-134.

[20]Hauser, J.R., Urban, G.L., Liberali, G., and Braun, M. 2009. Website morphing. Marketing Science, **28**(2): 202-223.

[21]Hayes, A.F. 2008. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach: Guilford Press.

[22]Heitmann, B., Kim, J.G., Passant, A., Hayes, C., and Kim, H.-G. 2010. An architecture for privacy-enabled user profile portability on the Web of Data. In Proceedings of the 1st International Workshop on Information Heterogeneity and Fusion in Recommender Systems. 16-23.

[23]Kay, J. and Kummerfeld, B. 2006. Scrutability, user control and privacy for distributed personalization. In Proceedings of the CHI 2006 Workshop on Privacy-Enhanced Personalization. 21-22.

[24]Kay, J., Kummerfeld, B., and Lauder, P. 2002. Personis: a server for user models. In Proceedings of the Adaptive Hypermedia and Adaptive Web-Based Systems. 203-212.

[25]Kim, H.-W., Chan, H.C., and Gupta, S. 2007. Value-based adoption of mobile internet: an empirical investigation. Decision Support Systems, **43**(1): 111-126.

[26]Kittur, A., Chi, E.H., and Suh, B. 2008. Crowdsourcing user studies with Mechanical Turk. In Proceedings of the Conference on Human Factors in Computing Systems 2008. 453-456.

[27]Knijnenburg, B.P. and Kobsa, A. 2013. Making decisions about privacy: information disclosure in context-aware recommender systems. ACM Transactions on Interactive Intelligent Systems (TiiS), **3**(3): 20.

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

[29]Komiak, S.Y. and Benbasat, I. 2006. The effects of personalization and familiarity on trust and adoption of recommendation agents. MIS quarterly: 941-960.

[30]Koren, Y., Bell, R., and Volinsky, C. 2009. Matrix factorization techniques for recommender systems. Computer, **42**(8): 30-37.

[31]Lampe, C., Ellison, N.B., and Steinfield, C. 2008. Changes in use and perception of Facebook. In Proceedings of the 2008 ACM conference on Computer supported cooperative work. 721-730.

[32]Li, T. and Unger, T. 2012. Willing to pay for quality personalization&quest; Trade-off between quality and privacy. European Journal of Information Systems, **21**(6): 621-642.

[33]Lin, J., Amini, S., Hong, J.I., Sadeh, N., Lindqvist, J., and Zhang, J. 2012. Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 501-510.

[34]Malhotra, N.K., Kim, S.S., and Agarwal, J. 2004. Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. Information Systems Research, **15**(4): 336-355.

[35]Margulis, S.T. 2003. Privacy as a social issue and behavioral concept. Journal of social issues, **59**(2): 243-261.

[36]Metzger, M.J. 2006. Effects of site, vendor, and consumer characteristics on web site trust and disclosure. Communication Research, **33**(3): 155-179.

[37]Milne, G.R. and Boza, M.-E. 1999. Trust and concern in consumers’ perceptions of marketing information management practices. Journal of Interactive marketing, **13**(1): 5-24.

[38]Mount, M.K., Barrick, M.R., Scullen, S.M., and Rounds, J. 2005. Higher - order dimensions of the big five personality traits and the big six vocational interest types. Personnel Psychology, **58**(2): 447-478.

[39]Peppers, D., Rogers, M., and Dorf, B. 1999. Is your company ready for one-to-one marketing. Harvard Business Review, **77**(1): 151-160.

[40]Pommeranz, A., Broekens, J., Wiggers, P., Brinkman, W.-P., and Jonker, C.M. 2012. Designing interfaces for explicit preference elicitation: a user-centered investigation of preference representation and elicitation process. User Modeling and User-Adapted Interaction, **22**(4-5): 357-397.

[41]Sheehan, K.B. and Hoy, M.G. 2000. Dimensions of privacy concern among online consumers. Journal of Public Policy & Marketing, **19**(1): 62-73.

[42]Shneiderman, B. 2000. Designing trust into online experiences. Communications of the ACM, **43**(12): 57-59.

[43]Smith, H.J., Milberg, S.J., and Burke, S.J. 1996. Information Privacy: Measuring Individuals' Concerns About Organizational Practices. MIS quarterly, **20**(2).

[44]Sundar, S. and Marathe, S. 2006. Is it tailoring or is it agency? Unpacking the psychological appeal of customized news. In Proceedings of the the 89th Annual Convention of the Association for Education in Journalism and Mass Communication.

[45]Sundar, S.S., Kang, H., Wu, M., Go, E., and Zhang, B. 2013. Unlocking the privacy paradox: do cognitive heuristics hold the key? In Proceedings of the CHI'13 Extended Abstracts on Human Factors in Computing Systems. 811-816.

[46]Toch, E., Wang, Y., and Cranor, L.F. 2012. Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. User Modeling and User-Adapted Interaction, **22**(1-2): 203-220.

[47]Tsai, J.Y., Egelman, S., Cranor, L., and Acquisti, A. 2011. The effect of online privacy information on purchasing behavior: An experimental study. Information Systems Research, **22**(2): 254-268.

[48]Unni, R. and Harmon, R. 2007. Perceived effectiveness of push vs. pull mobile location based advertising. Journal of Interactive advertising, **7**(2): 28-40.

[49]Westin, A.F. 1968. Privacy and freedom. Washington and Lee Law Review, **25**(1): 166.

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

[51]Xu, H., Dinev, T., Smith, J., and Hart, P. 2011. Information Privacy Concerns: Linking Individual Perceptions with Institutional Privacy Assurances. Journal of the Association for Information Systems, **12**(12).

[52]Xu, H., Luo, X.R., Carroll, J.M., and Rosson, M.B. 2011. The personalization privacy paradox: an exploratory study of decision making process for location-aware marketing. Decision Support Systems, **51**(1): 42-52.

[53]Xu, H., Teo, H.-H., Tan, B.C., and Agarwal, R. 2012. Research Note-Effects of Individual Self-Protection, Industry Self-Regulation, and Government Regulation on Privacy Concerns: A Study of Location-Based Services. Information Systems Research, **23**(4): 1342-1363.

[54]Zhang, B., Wu, M., Kang, H., Go, E., and Sundar, S.S. 2014. Effects of security warnings and instant gratification cues on attitudes toward mobile websites. In Proceedings of the 32nd annual ACM conference on Human factors in computing systems. 111-114.

7. APPENDIX

A. Scenarios

Condition			Scenario
Without Control	Demographics		How **CONCERNED** would you feel if Amazon.com accessed the following information about you in return for personalized recommendations, without asking you first?
	Products	Implicit Input	Suppose **YOU HAVE PURCHASED** the following items from Amazon.com. Please indicate how **CONCERNED** you would feel for Amazon.com to access your purchase history of each of the following items in return for personalized recommendations.
		Explicit Input	Suppose you **HAVE PURCHASED** the following items from Amazon.com. Please indicate how **CONCERNED** you would feel to provide your **RATINGS** of the items to Amazon.com in return for personalized recommendations.
With Control	Demographics		Suppose you **HAVE CONTROL** over the extent to which Amazon.com can access your personal information. With such control, how **CONCERNED** would you feel having the following information about you stored on Amazon.com?
	Products	Implicit Input	Suppose you **HAVE CONTROL** over the extent to which the following items **IN YOUR PURCHASE HISTORY** can be accessed by Amazon.com. With such control, please indicate how **CONCERNED** you would feel having each of the items in your purchase history on Amazon.com.
		Explicit Input	Suppose you **HAVE CONTROL** over the extent to which **YOUR RATINGS** of the following items can be accessed by Amazon.com. With such control, please indicate how **CONCERNED** you would feel to RATE each of the items in return for personalized recommendations.

B. Information Items

Related questions (see Appendix A) were answered on a scale of “1 = Not Concerned at All” to “7 = Extremely Concerned.”

Demographic Information	Product-Related Information
1. Gender	1. Textbooks
2. Age	2. Digital Games
3. Education	3. Jewelry
4. Race	4. Furniture
5. Relationship status	5. Snack Food
6. Technology use	6. Flowers
7. Email address	7. Shoes
8. Phone number	8. Laptop
9. Credit card number	9. Lingerie
10. Social security number	10. Condoms
11. Date of birth	11. Lubricant
12. Name	12. Book – Depression
13. Home address	13. Weight Loss Products
14. Company	14. Pregnancy Test
15. Interest areas	15. Book – Bankruptcy
16. Field of work	16. Fertilizer
17. Household income	17. Adult Diapers
18. Location	18. Hunting Knife
19. Calendar data	19. Cigarettes
20. Web browsing history	20. Bottle of Peroxide
21. IP address	21. Sex Toys
	22. HIV Test
	23. Pornographic DVD
	24. STD Medication
	25. Bulletproof Jacket
	26. Book - Bomb-Making

C. Measurements

These measures were all based on a scale of “1 = Strongly Disagree” to “7 = Strongly Agree” unless otherwise noted.

Perceived Value of Information Disclosure

1. I think my benefits gained from using Amazon.com’s service can offset the risks of my information disclosure.
2. The value I gain from using Amazon.com’s service is worth the information I give away.
3. I think the risks of my information disclosure will be greater than the benefits gained from using Amazon.com’s service.

Trust **(**Please indicate how well each of the following adjectives describes Amazon.com.)

1. Reliable
2. Trustworthy
3. Dependable
4. Honest
5. Fair
6. Exploitative (reverse coded)

Perceived Value of Online Personalization

1. I value web pages that are personalized for the device (e.g., computer, tablet, mobile phone, etc.), browser (e.g., Internet Explorer, Firefox, Chrome, etc.) and operating system (e.g. Windows, Mac OS, Unix) that I use.
2. I value websites that are personalized for my usage experience preferences.
3. I value websites that acquire my personal preferences and personalize the services and products themselves.
4. I value goods and services that are personalized based on information that is collected automatically (e.g., IP address, web browsing history) but cannot identify me as an individual.
5. I value goods and services that are personalized based on information that I have voluntarily given out (e.g., age, household income, field of work) but cannot identify me as an individual.
6. I value goods and services that are personalized on information I have voluntarily given out and can identify me as an individual (e.g., name, address, credit card number).

Importance of Control

1. It is important for me to restrict Amazon.com’s use of a specific type of information for personalized recommendations.
2. It is important for me to control Amazon.com’s access of a specific type of information for personalized recommendations.
3. It is important for me to control the amount of information accessed by Amazon.com for personalized recommendations.

General Privacy Concern

1. I am sensitive about giving out information regarding my preferences.
2. I am concerned about anonymous information (information collected automatically but cannot be used to identify me, such as my computer, network information, operating system, etc.) that is collected about me.
3. I am concerned about how my personally un-identifiable information (information that I have voluntarily given out but cannot be used to identify me, e.g., age, gender, field of work, etc.) will be used by firms.
4. I am concerned about how my personally identifiable information (information that I have voluntarily given out AND can be used to identify me as an individual, e.g., name, home address, credit card number, etc.) will be used by firms.

Demographics (These were all posed as multiple-choice questions)

1. What is your age?
2. What is your gender?
3. What was the highest level of education you have received?
4. What racial group do you belong to?