Journal of Advertising

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ujoa20

When E-Commerce Personalization Systems Show and Tell: Investigating the Relative Persuasive Appeal of Content-Based versus Collaborative Filtering

Mengqi Liao & S. Shyam Sundar

JOURNAL OF ADVERTISING
2022, VOL. 51, NO. 2, 256–267
https://doi.org/10.1080/00913367.2021.1887013

RESEARCH NOTE

When E-Commerce Personalization Systems Show and Tell: Investigating the Relative Persuasive Appeal of Content-Based versus Collaborative Filtering

Mengqi Liao and S. Shyam Sundar

Pennsylvania State University, University Park, Pennsylvania, USA

ABSTRACT
In the e-commerce context, are we persuaded more by a product recommendation that matches our preferences (content filtering) or by one that is endorsed by others like us (collaborative filtering)? We addressed this question by conceptualizing these two filtering types as cues that trigger cognitive heuristics (mental shortcuts), following the heuristic-systematic model in social psychology. In addition, we investigated whether the degree to which the recommendation matches user preferences (or other users' endorsements) provides an argument for systematic processing, especially for those who need deeper insights into the accuracy of the algorithm, particularly in product categories where quality is subjective. Data from a 2 (algorithm type: content vs. collaborative filtering) x 3 (percentage match: low vs. medium vs. high) x 2 (product category: search vs. experience) + 2 (control: search and experience) between-subjects experiment ($N = 469$) reveal that for experience products, consumers prefer content-based filtering with higher percentage matches, because it is perceived as offering more transparency. This is especially true for individuals with high need for cognition. For search products, however, collaborative filtering leads to more positive evaluations by triggering the "bandwagon effect." These findings have implications for theory pertaining to the use of artificial intelligence in strategic communications and design of algorithms for e-commerce recommender systems.

The vast majority of consumers prefer tailored product offerings, with personalization positively predicting purchase likelihood (Epsilon Marketing 2018). When e-commerce systems offer a personalized recommendation, they often provide a brief rationale for suggesting a product. In some cases, it is based on past purchase history or similarity to the products liked before. In others, it is based on preferences or behaviors of similar users. These two types of explanations refer to two different types of filtering algorithms—content-based and collaborative filtering respectively—that have been widely adopted in the industry (Pearl, Chen, and Rong 2012), but it is unclear whether they elicit different psychological responses from users.

Unlike content-based recommender systems, collaborative systems leverage the power of a larger group of similar users and thus are more likely to trigger the "bandwagon heuristic" (the notion that "if others think that something is good, then I should think so too"; Sundar 2008, p. 83). Could this be more persuasive than simply relying on one's own preferences? Or does it depend on the nature of products? Would a consumer prefer content filtering when s/he is shopping for products that are *searchable* (i.e., most of the products' information could be obtained easily before the purchase) and collaborative filtering when considering products that are *experiential* (i.e., products for which full information is difficult to acquire in advance)? Or does it depend on the strength of the recommendations (Herlocker et al. 2004), that is, the predicted accuracy or the extent to which the system thinks the users would like the recommended products (% match)?

Some systems indicate the extent to which the recommended items match with one's history (e.g., Netflix's percentage matching) or those of others (e.g., Amazon's common statement "a percentage of people who bought this also bought that"). Unlike superficial *cues* that trigger cognitive heuristics (like the

CONTACT Mengqi Liao mul914@psu.edu Penn State University Park, Donald P. Bellisario College of Communications, University Park, PA 16802-5101, USA.
Mengqi Liao (MA, Pennsylvania State University) is a PhD student, Donald P. Bellisario College of Communications, Pennsylvania State University.
S. Shyam Sundar (PhD, Stanford University) is James P. Jimirro Professor of Media Effects and Co-director of the Media Effects Research Laboratory, Donald P. Bellisario College of Communications, Pennsylvania State University.

aforementioned bandwagon heuristic), such indicators of accuracy and recommendation strength constitute *arguments* that may promote systematic, rather than heuristic, processing (Chaiken 1980), particularly among consumers who care deeply about the product or those who have a higher innate need to know.

The current study aims to explore the psychological factors underlying consumers' experience of recommender systems by examining the interactive effects of system characteristics, product types, and user traits. This is important not only because we will learn more about how different recommendation algorithms differentially affect user trust but also because this knowledge can inform the design of future AI systems that lead to better e-commerce experiences.

Literature Review

First-generation recommender systems use content-based filtering by matching users' preferences with specific features of the products (Burke 2007). This approach focuses on each user's own preference (Degemmis et al. 2004). Second-generation recommender systems use collaborative filtering and are more socially oriented (Ochi et al. 2010). They not only take into account users' personal preferences of specific product characteristics but also recommend products that are consumed or endorsed by users who share some similarities with the recipients of the recommendations (Pearl, Chen, and Rong 2012).

While research has focused on the relative performance of these two algorithms and how they could be fine-tuned and advanced to produce more accurate matches (Ochi et al. 2010), we do not know much about how users' subjective perceptions toward these two types of algorithms influence their evaluation of recommendation systems. The logic of a specific algorithm used by a recommendation system is usually explained to users in one or two sentences on the interface. Many content-based filtering recommendation systems, for example, say something like "*Based on your past selections, you might be interested in …*" while those for collaborative filtering say "*Other users with similar interests also enjoyed …*".

Theoretically, both these explanations can be seen as persuasive appeals by offering important cues for decision-making. Cues for efficient decision-making have long been an object of study in social psychology and communication. The heuristic-systematic model (HSM) was among the earliest theoretical formulations to highlight the power of cues by positing that individuals have an innate preference to make decisions based on superficial cues because they trigger cognitive heuristics, or mental shortcuts, for quick decision-making, thereby eschewing the need for effortful consideration of underlying information. Chaiken (1980) labeled this "heuristic processing" and contrasted it with the more effortful "systematic processing." HSM proposes that group endorsement could serve as a simple cue that persuades individuals by triggering the heuristic that "consensus implies correctness" (Chaiken, Akiva, and Eagly 1989, p. 216). Applying the notion of consensus effects to users' interaction with digital media interfaces, Sundar (2008) proposed the *bandwagon heuristic*, or the notion that "if others think that something is good, then I should think so too" (Sundar 2008, 83), which could aid consumers' online decision-making. Several studies have shown that the bandwagon heuristic could be influential in individuals' credibility assessment of online information (Sundar and Nass 2001) and promoting subsequent purchasing behaviors (Sundar, Oeldorf-Hirsch, and Xu 2008). This suggests a theoretical mechanism for the relative popularity of collaborative filtering in the e-commerce marketplace. Thus, we propose the following hypothesis:

> **H1:** Collaborative filtering will lead to higher level of bandwagon perception among users, which will be associated with a more positive evaluation of the recommendation system, compared to content-based filtering.

Predictive Accuracy of Algorithm

While disclosure of the type of algorithm used (collaborative vs. content) may influence users superficially by simply providing the source of information, indicators of the quality of resulting recommendations will likely require deeper processing, calling for a more systematic consideration of algorithm performance, especially in terms of its prediction ability (Herlocker et al. 2004; Adomavicius and Tuzhilin 2005). Accuracy is by far the most important criterion for assessing the quality of recommendation systems in the industry (Zhang and Xu 2018) and, as such, constitutes an argument that is processed more systematically than the global nature of the algorithm.

One way to improve consumers' evaluations is to inform them about objective accuracy of the recommendation system (Zhang and Xu 2018; Knijnenburg et al. 2012). A common external indicator of accuracy is percentage match (% match). Netflix, for example, shows "xx% match" to every show or movie it recommends. According to Netflix, it is calculated based on preferences of similar users, that is, users who share the same streaming history (Netflix 2020). Google Maps, on the other hand, recommends restaurants based on the user's own past ratings, location history, and the match of that user's food and drink preferences with information from the restaurant (Debczak 2018). Recently, an app called Taste was launched in the hope of creating taste profiles for its users such that each individual can get a personalized "match %" for a variety of products (Smith 2019).

As posited by the additivity principle of HSM, high percentage matching might be an important amplifier of the bandwagon perception when users are interacting with collaborative filtering recommendation systems, as a heightened crowd of similar users is made salient on the interface. Research on the effects of interface nudges has found that if a personalized energy recommender interface employs the similarity social norm (e.g., xx% of customers similar to you do this to save energy) with a higher (75%) rather than lower percentage (18%), users are more likely to conform to the norm of the similar users and eventually be persuaded to adopt the advocated behaviors, such as lowering the thermostat to 14 °C before going to bed (Starke, Willemsen, and Snijders 2020). This suggests that bandwagon perception will be more pronounced in the presence of high percentage match, leading us to propose the following moderated mediation hypothesis:

> **H2:** The indirect effects of recommendation system algorithm type on users' evaluation of the recommendation system through bandwagon perception will be higher when the percentage match is high rather than low.

A high percentage match not only provides additional persuasive information as per additivity principle but also signals greater transparency by revealing why the system made the recommendation (Biran and Cotton 2017). When the system justifies whether the products are recommended based on similar users' preferences (collaborative filtering) or based on the user's own preference (content-based filtering), it is offering fundamentally different explanations of its underlying mechanism. It is possible that users find greater certainty in knowing that the system personalizes based on one's own preferences rather than those of unknown others, thereby leading to greater perceived transparency in content-based filtering. But this is true only under conditions of high percentage match because the recommendations clearly reflect one's preferences. However, when recommendations do not cohere with one's prior choices, collaborative filtering may seem like a more satisfactory explanation for the system's performance because of the user's lack of awareness of others' preferences and choices. Therefore, the degree to which users perceive transparency in the system may depend on the extent to which they have access to corroborating evidence about preferences—since users know their own preferences, they can easily judge the veracity of a high versus low percentage match in the case of content-based filtering, but they cannot be so certain in the case of collaborative filtering. A system's explainability has long been considered an important contributor of user trust, defined as attitudes toward automated systems that affect users' reliance on the system (Lee and See 2004). Perceived transparency is found by many researchers as facilitating trust (Chen and Sundar 2018; Eslami et al. 2018). Thus, when a system touts a high percentage match, it conveys greater transparency beyond revealing the filtering method. After all, the system should only recommend items to users that it thinks the users will like. But, when a recommendation is accompanied by a low percentage match, it will beg the question: why then did the system recommend the product? Users will be left to surmise that there could be other factors involved in providing the recommendation—factors that are not made apparent to users, thus leaving them wanting in terms of transparency and thereby leading to attenuation of the cue effect of algorithm type.

Therefore, we propose the following hypothesis:

> **H3:** The indirect effects of recommendation system algorithm type on users' evaluation of the recommendation system through perceived system transparency will be higher when the percentage match is high rather than low.

However, as revealed by past literature on exemplification (Zillmann and Hans-Bernd 2000), such numerical indicators of prior probability of outcomes as percentage match, also called base-rates, are often overlooked by individuals because they are inherently difficult to process, given their pallid nature, especially when compared to vivid descriptions of exemplars. Thus, as argued by HSM, only certain individuals with cognitive ability and motivation are likely to factor such base-rate information into their perceptions. This calls for a consideration of individual differences in processing the cognitively effortful arguments pertaining to performance of recommendation system algorithms.

Individual Differences

In the dual process literature, need for cognition (NFC) is a stable individual difference that often moderates how individuals react to persuasive appeals. Consumers who are high in NFC engage in more effortful thinking when making purchasing decisions, compared to those with low NFC (Cacioppo and Petty 1982; see Petty, and Evans 2009). Consistent with the propositions in HSM, individuals who are high in need for cognition will be more influenced by the quality of the persuasive messages (Lee and Yoon Jae 2010) and rely less on mental shortcuts (Cacioppo et al. 1986), like the bandwagon heuristic mentioned previously. Therefore, we propose the following hypothesis:

H4: The effects of recommendation system algorithm type on users' evaluation of the recommendation system will be higher for low-NFC individuals compared to high-NFC individuals.

Past research has also revealed that individuals high in NFC will be more likely to gather relevant information before making a purchase decision on a shopping website (Lee and Yoon Jae 2010); thus, the percentage matching information might be valued more by individuals high in NFC. Based on this rationale, we propose the following hypothesis:

H5: For individuals high in NFC, there will be an interaction between recommendation system algorithm type and percentage matching on users' evaluation of the recommendation system.

Method

A 2 (algorithm type: content vs. collaborative filtering) x 3 (percentage match: low vs. medium vs. high) x 2 (product category: search vs. experience) + 2 (control for search and experience) between-subjects online experiment was conducted to test the proposed hypotheses. In addition to the theorized independent variables of algorithm type and percentage match, we included product category as a variable because it is known that a content-based recommender system is generally preferred by users when products are *searchable* (e.g., appliances, furniture, electronics), namely products for which most of the information can be obtained easily before the purchase, whereas collaborative filtering is preferred when the products are *experience* products (e.g., books, food, travel, beauty salon) because comprehensive information is difficult to obtain in advance (Aggarwal and Vaidyanathan 2005). With this in mind, we created an experience product e-commerce website (a travel destination recommendation system called Wheretogo.com) and a search product site (a smartwatch recommendation system called shopbotics.com) especially for this study.

Procedure

Participants were first invited to fill out a pre-questionnaire through which their demographic information and levels of NFC were obtained. They were then randomly assigned to one of the experimental conditions and directed to the corresponding mock website.

On the website, they were first asked to answer a series of questions to build their user profile. These questions were relevant to the specific product recommended, but not so specific that they would make the subsequent recommendations appear as obviously right or wrong. Sample profile-building questions for travel destinations recommendation system include "Do you like to try new things?" and "Would it be a challenge for you to spend the whole weekend all by yourself without feeling bored?" Sample profile-building questions for smartwatch recommendation system include "Do you often make decisions on a whim?" and "How long do you plan on your average exercise activity lasting?" After their profiles were successfully built, they were presented with a series of recommended items with the specific manipulation cues on the interface (described in the following section). After interacting with the recommendation system for at least 3 minutes, they were directed back to the questionnaire, which elicited their evaluations of the recommendation system.

Participants

A total of 586 participants were recruited through Amazon Mechanical Turk (MTurk). After eliminating those who did not pass the two attention checks, final data from 469 participants were included in the analysis, with 394 participants in experimental conditions and the remaining 75 in the two control conditions.

Among the participants, 56.5% ($n = 265$) were male, 42.6% were female ($n = 200$), 0.6% ($n = 3$) identified their gender as other or not listed, and 0.2% ($n = 1$) preferred not to answer, with an average age of 33.99 ($SD = 10.49$, range = 18–74). Our participants were relatively highly educated, with 79.7% ($n = 372$) having a bachelor's degree or above.

A power analysis was conducted using an alpha of 0.05, a power of 0.8, and a medium effect size ($f = 0.25$) (Faul et al. 2007). Excluding the two control conditions (so that we can run balanced, full-factorial analyses for directly addressing the hypotheses and

Figure 1. a. Landing page and result directory page for content-based recommendation systems. b. Landing page and result directory page for collaborative recommendation systems.

research questions), we have 12 conditions and 8 degrees of freedom, for which the desired sample size is 372, which is smaller than our sample size of 394, indicating sufficient statistical power for our analysis.

Manipulation

The recommended items provided to the participants were held constant across conditions. Only the specific cues that signal the algorithm types adopted by the recommendation system as well as the percentage matching were manipulated. The welcome page and profile-building page were designed to be the same. After the profile building, when they clicked the next button, participants were presented with a landing page and a results directory page. Figure 1a shows the interface individuals saw when interacting with the content-based recommendation system, whereas Figure 1b shows the interface of the collaborative recommender system.

To eliminate the confounding effects of price, all recommendations included two low-budget, two medium-budget, and two high-budget options, that is, a total of six products within the same product type. On every page of the search results, content-based recommendation system included the sentence *"Based on your own preferences, you might like … "* and collaborative ones said *"Based on preferences from similar users, you might like … "* (see Figure 1a and 1b).

Next to each of the recommended products, the manipulation of percentage match appeared in the format of both a number and a progress bar next to the recommended products (see Figure 2a and 2b). Individuals assigned to content-based recommendation systems with high percentage match saw "92% (96%, 89%, 94%, 95%, 90%) match to your preferred characteristics," while those assigned to the medium percentage matching saw "77%, 81%, 74%, 79%, 80%, 75%" and those assigned to the low percentage matching saw "62%, 66%, 59%, 64%, 65%, 60%". Users exposed to collaborative recommendation systems saw the same number for high (92%, 96%, 89%, 94%, 95%, 90%), medium (77%, 81%, 74%, 79%, 80%, 75%"), and low (62%, 66%, 59%, 64%, 65%, 60%) percentage matching, but the wording was changed to *"xx% of users similar to you used it/traveled there."* All participants could click the button *"View more"* to see a short description of all of the recommended items, and these descriptions were held constant across conditions.

Measurement

All of the variables were measured using 7-point Likert scales, with participants being asked to rate a series of statements.

Control Variable

Need for uniqueness (NFU), or the tendency or need to be able to differentiate oneself from others and be treated as one of a kind (Snyder and Fromkin 1977), might override the effects of recommendation system types on user evaluation, and was therefore controlled in our study. It was measured through the Self-Attributed Need for Uniqueness Scale developed by

Lynn and Judy (1997). The scale contained 4 items, e.g., *"I prefer being different from other people."* The overall reliability of this scale was acceptable (α = 0.87, *M* = 5.16, *SD* = 1.13).

Moderating Variable

NFC was measured through the short-form 5-item Need for Cognition scale (Coelho, Hanel, and Wolf 2020). Sample questions include *"I would prefer complex to simple problems"* and *"I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought"* (α = 0.77, *M* = 5.17, *SD* = 1.13).

Manipulation Check

The manipulation check item for content-based recommendation system was *"The system recommends items based on matching their characteristics with my preferences,"* and that for the collaborative system was *"The system recommended the above travel destinations/products because similar users travelled there/used them."* The manipulation check for percentage matching in the content-based conditions was *"The recommended travel destinations/products are highly matched to my preferences,"* while that for the collaborative conditions was *"A high percentage of similar users have travelled to/used the recommended travel destinations/products."*

Figure 2. a. Percentage matching for experience product (content-based vs. collaborative). b. Percentage matching for search product (content-based vs. collaborative).

Mediating Variables

Bandwagon perception (BP) was measured with 5 items adopted from past research (Sundar, Oeldorf-Hirsch, and Xu 2008). Example items include *"Many users chose the suggested destinations/products above"* and *"Many users seem to be using this system"* (α = 0.87, *M* = 5.06, *SD* = 1.07).

Perceived transparency was measured with 2 items (Knijnenburg et al. 2012) including *"I understood why the travel destinations/smartwatches were recommended to me"* and *"The system helps me understand why the travel destinations/smartwatches were recommended to me"* (α = 0.80, *M* = 5.11, *SD* = 1.03).

Dependent Variables

Purchase intention was measured with 3 items including *"I intend to visit/buy (at least one of) the suggested destinations/products"* and *"I will follow up and look for information about (at least one of) the suggested destinations/products in other websites"* (α = 0.86, *M* = 5.14, *SD* = 1.41).

Perceived recommendation system quality (RS quality) was measured with 6 items from the scale developed by Knijnenburg and colleagues (Knijnenburg et al. 2012). Sample questions include *"The destinations (items) recommended to me matched my interests well"* and *"The recommender is useful"* (α = 0.87, *M* = 5.14, *SD* = 1.41).

Perceived trust toward the recommendation system (RS trust) was measured by adopting the 2-item scale developed by Hsiao et al. (2010): *"I think that the travel destination (product) recommendations of this* *system are credible*" and "*I trust this recommendation system*" ($\alpha = 0.79$, $M = 5.18$, $SD = 1.19$).

Data Analysis

The Process macro in SPSS (Hayes 2018) was used to test the hypotheses. Model 4 was used for H1. Model 7 was used for H2 and H3. Model 7 was used for testing H4, and Model 12 was used to test H5. All of the hypothesis testing using Process macro were run three times: combined data (search and experience products combined), search products only, and experience products only.

Results

Manipulation Check

A one-way analysis of variance (ANOVA) indicated that when the participants were presented with content-based filtering, they perceived the system to utilize the content-based filtering rules ($M = 5.52$, $SD = 1.15$) more than participants exposed to collaborative filtering ($M = 5.28$, $SD = 1.22$, $p < .05$) and those in the control conditions ($M = 4.93$, $SD = 1.47$, $p <. 05$), $F(2, 466) = 6.48$, $p < .01$. Similarly, when participants were presented with collaborative filtering ($M = 5.56$, $SD = .97$), they perceived the system to utilize the collaborative filtering rules significantly more than participants in content-based filtering ($M = 5.09$, $SD = 1.28$, $p < .001$) and the control conditions ($M = 5.14$, $SD = 1.08$, $p < .01$), $F(2, 466) = 9.24$, $p < .001$.

One-way ANOVAs showed that for participants exposed to content-based filtering, the higher the percentage shown on the interface, the higher the perceived match of the products'/travel destinations' characteristics to their own preferences (low percentage: $M = 4.97$, $SD = 1.34$; medium percentage: $M = 5.26$, $SD = 1.39$; high percentage: $M = 5.60$, $SD=.98$), $F(2, 199) = 3.51$, $p = .03$. Another one-way ANOVA for participants exposed to collaborative filtering also showed that the mean for the high-percentage condition was the highest ($M = 5.52$, $SD = 1.06$), followed the medium-percentage ($M = 5.40$, $SD = 1.12$), and the low-percentage conditions ($M = 5.38$, $SD = 1.00$), although these mean differences were not statistically significant, $F(2, 189) = .32$, $p = .70$. A pretest with 190 participants revealed that the percentage matching manipulation for collaborative filtering conditions approached significance, $F(2, 82) = 5.85$, $p = .09$, with high percentage matching being perceived as high ($M = 5.81$, $SD = .80$), followed by medium percentage matching ($M = 5.41$, $SD = .95$) and low percentage matching ($M = 5.15$, $SD = 1.46$).

Mediating Effects of Bandwagon and Perceived Transparency

First, to test H1, Model 4 of the Process macro was utilized to test whether bandwagon perception mediates the relationship between recommendation system (RS) type and dependent variables for the combined data when the product type, percentage matching, and individuals' level of NFC and NFU are held constant. While no significant mediating effects were observed for the combined data, for search products, bandwagon perception was found to significantly mediate the effects of RS type on purchase intention, perceived RS quality, and trust (see Table 1). Specifically, collaborative filtering was associated with higher bandwagon perception ($M = 5.36$, $SD = .89$) compared to content-based filtering ($M = 5.11$, $SD = 1.02$) and in turn led to higher users' evaluations of the RS (see Table 3). Therefore, H1 was supported for search products. But analyses using Model 7 of the Process macro revealed no significant evidence of percentage match moderating the mediation of bandwagon perception, and no significant effects were observed. The same test was run to test H3 with transparency as the mediator. Results revealed no significant evidence of percentage match moderating the mediation of transparency for the combined data when individuals' level of NFC and product type were held constant or for search and experience products conditions separately. Thus, H2 and H3 were not supported.

The Moderating Effects of NFC

H4 predicted that NFC would moderate the effects of RS type on the dependent variables. Results from Model 7 showed that NFC did not significantly moderate the mediation effects of bandwagon perception on any of the dependent variables for either product type. Therefore, H4 was not supported.

H5 hypothesized that there would be a three-way interaction effect among percentage match, RS type, and NFC on users' evaluation of the system. Model 12 was adopted to test whether there is an indirect three-way interaction effect (among RS type, percentage matching, and NFC) on the dependent variables either through bandwagon perception or perceived transparency. Results revealed no such three-way interaction effects for the combined data. However, for experience products, a moderated mediation effect of NFC was

Table 1. Summary of results from moderated mediation analysis for search product.

Mediation path	B^a	SE	CI[b] Lower CI	CI[b] Upper CI
RS type[c] → bandwagon perception → purchase intention*	.11	.05	.03	.22
RS type[c] → bandwagon perception → RS quality*	.10	.04	.02	.18
RS type[c] → bandwagon perception → RS trust*	.09	.04	.02	.18

[a]Unstandardized path coefficient.
[b]Bias-correlated and accelerated 95% confidence interval (CI).
[c]Recommendation system (RS) type is coded as 1 = content-based filtering, 2 = collaborative filtering
*$p < .05$

observed on transparency, which was positively associated with purchase intention, RS quality, and trust (see Table 2 for a summary of results). Specifically, content-based and collaborative based filtering had different effects, but only when the percentage match was high. Individuals low in need for cognition prefer a system that recommends products highly endorsed by similar users (collaborative filtering) while those high in NFC prefer recommendations that are highly matched to their own characteristics (content-based filtering) (see Figure 3). No such effects were observed for search products. When a similar test was run for bandwagon perception, no three-way interaction effects were observed. Therefore, H5 was supported for the mediating effect of transparency and for experience products.

Summary of Results

In sum, for search products, we found that consumers prefer a collaborative filtering system over a content-based recommendation system because it triggers the bandwagon heuristic, which positively influences user experience of the e-commerce website and promotes purchase intention.

For experience products, which are ambiguous by nature, individuals high in NFC prefer recommendations that match their preferences and characteristics (i.e., content-based filtering), whereas those low in NFC prefer recommendations that are endorsed by a majority of similar users (collaborative filtering), as they believe it offers more transparency.

Discussion

The current study found that different types of recommendation systems not only offer different ways for consumers to get product recommendations but also serve as cues that have significant persuasive appeal upon consumers' purchase intentions of the recommended products and their subjective evaluations of the recommendation system. Even though participants get the same product recommendations, exposure to cues on the interface indicating whether the product is recommended based on content-based or collaborative filtering and exposure to arguments relating to recommendation quality (i.e., its match with preferences) together influence their user experience. Therefore, a clear implication for practitioners is to focus more design energy on interface cues and arguments that communicate the nature and quality of the recommendation algorithm instead of focusing solely on perfecting the precision of algorithm predictions.

Contrary to what was found by Ochi et al. (2010), our data show that collaborative filtering is not optimal for experience products, nor is content-based filtering for search products. In fact, it is quite the opposite. Consumers prefer collaborative filtering over content-based filtering when they shop for products that have full information on the website (search products) because it triggers the bandwagon heuristic. Trust in others' opinions seems to have utility because it can save search costs involved in researching product information, looking up details, analyzing them, and picking out the right product. Since much of the information about search products is public, consumers are quite comfortable in letting others sift through the information and describe their experience with the products.

Collaborative filtering with a high percentage match appears to work better for experience products as well, but only for individuals with low NFC. Those high in NFC prefer content-based filtering with high percentage match because they perceive it as being more transparent. Indeed, system transparency might be valued more when it recommends products for which full information is difficult to obtain prior to purchase (experience products), as these are known to cause uncertainty among consumers (Weinschenk 2009). Thus, for experience products, consumers might be more inclined to collect information from the e-commerce environment to reduce their uncertainty about how the products came to be recommended, including the types of the recommender system and percentage match. In general, we found that percentage

Table 2. Summary of results from moderated mediation analysis for experience product.

Mediation path	B[a]	SE	CI[b] Lower CI	Upper CI
RS type[d]*percentage*nfc[c]→ transparency→ purchase intention*	−.13	.07	−.29	−.01
RS type[d]*percentage*nfc[c]→ transparency→ RS quality*	−.16	.08	−.32	−.01
RS type[d]*percentage*nfc[c]→ transparency→ RS trust*	−.18	.09	−.36	−.02

[a]Unstandardized path coefficient.
[b]Bias-correlated and accelerated 95% confidence interval (CI).
[c]nfc: need for cognition.
[d]Recommendation system (RS) type is coded as 1 = content-based filtering, 2 = collaborative filtering
*$p < .05$.

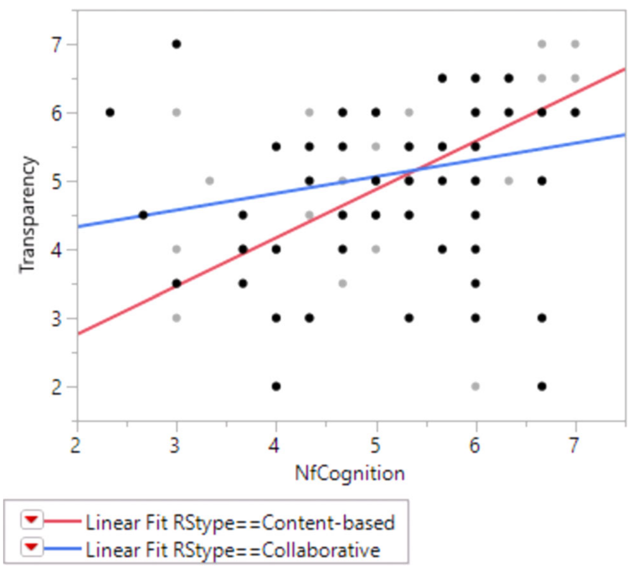

Figure 3. Moderated effects of need for cognition on perceived transparency for experience products when percentage match is high.

match is valued more for recommendations of experience products compared to search products, probably because it provides another important piece of information to reduce their uncertainty. This finding has implications for industry practitioners. There is a growing trend of attaching percentage match to recommended products. Netflix, for example, shows the "x% match" to every show or movie it recommends. Google Maps also launched an additional way of recommending restaurants by showing how well an individual "matches" with certain restaurants. Our findings suggest that resources required for compiling this indicator of accuracy are better spent on recommendations of experience products rather than search products. High percentage match scores might serve to reduce user uncertainty and help increase purchase intention of equivocal products or experiences (e.g., a movie, a travel experience, a dining experience). Moreover, sites and services would do well to clearly articulate the locus of these matches (either to the user's preferences, as in content-based filtering, or to other users' preferences, as in collaborative filtering) given their differential appeal to low-NFC and high-NFC consumers. If a site primarily serves high-NFC consumers, then high matches with one's personal preferences would be optimal, whereas if it serves a wider base, especially of low-NFC consumers and/or heuristic processors of content, then high matches with similar others' preferences would be more desirable.

In general, results from the current study show that individual differences are powerful in shaping consumers' experience with recommender systems, particularly in NFC. The moderating effect of NFC for experience products can be explained by research in social cognition showing that individuals high in NFC are more likely to engage in systematic information processing (Chaiken 1980), while those low in NFC tend to rely more on heuristic processing and avoid cognitively demanding activities (Cacioppo and Petty 1984). Clearly, content-based filtering is more consonant with systematic processing because of its emphasis on scrutinizing various characteristics of the actual products for a match with user preferences. Collaborative filtering, on the other hand, is more akin to heuristic (or superficial) processing in that it relies on communal endorsement, which is peripheral to product characteristics. It appears that low-NFC individuals find this method of determining recommendations to be easier to understand compared to content-based filtering (which would require more effort in order to think through the ways in which the system might convert their responses on the profile-building page to recommendations in seemingly unrelated domains such as travel destinations). Since ease of understanding was central to our measure of perceived transparency, it stands to reason that this variable mediated the effect of algorithm type differently for low-NFC and high-NFC individuals. Given the widespread prevalence of collaborative filtering, the low-NFC individuals find this logic for providing recommendations more accessible and therefore more transparent than the content-based logic, which requires them to imagine too much.

Fortunately, it is feasible these days for practitioners to gauge personality factors implicitly through

Table 3. Means and standard deviations for all evaluations of the recommendation system for experience and search products.

	Experience products						Search products					
	Collaborative		Content-based		Control		Collaborative		Content-based		Control	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Bandwagon perception	4.88	1.14	4.92	1.15	4.78	1.16	5.36	0.89	5.10	1.02	5.05	1.13
Perceived transparency	4.92	1.22	4.87	1.41	4.66	1.34	5.47	1.05	5.33	1.07	5.09	1.33
Purchase intention	5.02	1.48	5.23	1.46	5.16	1.42	5.11	1.32	5.19	1.40	4.85	1.51
RS quality	4.84	1.41	5.11	1.35	5.01	1.19	5.29	1.08	5.37	0.99	5.14	1.19
RS trust	5.01	1.22	5.12	1.35	4.86	1.29	5.27	1.05	5.29	1.11	5.08	1.33

RS: recommendation system.

history of users' interactions with recommender systems (Beheshti et al. 2020), mainly due to research that confirms a strong correlation between personality type and user preferences for products (Rentfrow, Goldberg, and Zilca 2011). In fact, such implicit methods of predicting users' personalities have increasingly been studied by recommender systems researchers (see Beheshti et al. 2020 for a review). The outcomes of their work can be aligned with our findings to achieve greater targeting precision and ensure optimal online shopping experiences for consumers.

Based on our results, practitioners could embed more collaborative filtering elements when recommending search products. This is already quite prevalent in e-commerce websites like Amazon, which show "customers who bought items in your shopping cart also bought xxx." For experience products, showing the higher percentage matching recommendations might be beneficial in increasing consumers' perceived transparency of the recommendation system, which could in turn lead to more positive persuasive outcomes. In addition, in the process of pursuing explainable artificial intelligence, our study discovered that simply telling consumers the type of algorithm the recommendations systems adopted with a single sentence on the interface can increase the system's transparency in users' minds. And, given the moderating effect of need for cognition, it is important for practitioners to pay greater attention to psychographics while designing algorithms and interfaces for future recommendation systems.

Limitations

The current study is not free from limitations. First, the manipulation of percentage match for collaborative filtering failed the manipulation check, in part because the questions we asked are perceptual rather than ontological. Participants might have different perceptions of what "a higher percentage of similar users" means—quite unlike in the content-based conditions, where the percentage of match between products characteristics and one's preference is more objective.[1] Furthermore, they may have processed that information nonconsciously or without storing it in memory. After all, it is a known fact that most consumers evaluate websites and take actions nonconsciously when online (Weinschenk 2009). More qualitative research is needed to understand how consumers react to these cues and arguments both consciously and nonconsciously.

Second, we only chose two products and the two most common types of recommendation systems. Future research would do well to investigate the psychological effects of other types of recommendation systems as well, and for a wider range of products, in order to verify the external validity of our findings.

Note

1. We ran a separate set of analyses treating the manipulation-check item (perceived percentage matching), rather than the manipulated percentage matching, as the independent variable and found the same results (for experience products that have high percentage match, individuals high in NFC prefer recommendations via content-based filtering rather than collaborative filtering, whereas those low in NFC prefer collaborative filtering over content-based filtering). As a result, we decided to retain the use of manipulated percentage matching as the independent variable, as that would afford us causal, rather than simply correlational, inferences.

References

Adomavicius, G., and A. Tuzhilin. 2005. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering* 17 (6):734–49. doi:10.1109/TKDE.2005.99

Aggarwal, P., and R. Vaidyanathan. 2005. Perceived effectiveness of recommendation agent routines: Search vs. experience goods. *International Journal of Internet Marketing and Advertising* 2 (1/2):38–55. doi:10.1504/IJIMA.2005.007503

Biran, O., and C. V. Cotton. 2017. Explanation and justification in machine learning: A survey. *IJCAI-17 Workshop on Explainable AI (XAI)* 8 (1):8.

Beheshti, A., S. Yakhchi, S. Mousaeirad, S. M. Ghafari, S. R. Goluguri, and M. A. Edrisi. 2020. Towards cognitive recommender systems. *Algorithms* 13 (8):176. doi:10.3390/a13080176

Burke, R. 2007. Hybrid web recommender systems. In *The adaptive web: Methods and strategies of web personalization*, eds. Peter Brusilovsky, Alfred Kobsa, and Wolfgang Nejdl, 377–408. Lecture Notes in Computer Science. Berlin: Springer.

Cacioppo, J. T., and R. E. Petty. 1982. The need for cognition. *Journal of Personality and Social Psychology* 42 (1): 116–31. doi:10.1037/0022-3514.42.1.116

Cacioppo, J. T., and R. E. Petty. 1984. The elaboration likelihood model of persuasion. In *NA - Advances in Consumer Research Volume* 11, ed. Thomas C. Kinnear, 673–675. Provo, UT: Association for Consumer Research.

Cacioppo, J. T., R. E. Petty, C. F. Kao, and R. Rodriguez. 1986. Central and peripheral routes to persuasion: An individual difference perspective. *Journal of Personality and Social Psychology* 51 (5):1032–43. doi:10.1037/0022-3514.51.5.1032

Chaiken, S. 1980. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology* 39 (5):752–66. doi:10.1037/0022-3514.39.5.752

Chaiken, S., L. Akiva, and A. H. Eagly. 1989. Heuristic and systematic information processing within and beyond the persuasion context. In *Unintended thought*, eds. J. S. Uleman and J. A. Bargh, 212–52. New York: The Guilford Press.

Chen, T.-W., and S. S. Sundar. 2018. This app would like to use your current location to better serve you: Importance of user assent and system transparency in personalized mobile services. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Paper no. 537, 1–13. New York: ACM.

Coelho, G. L. H., P. H. P. Hanel, and L. J. Wolf. 2020. The very efficient assessment of need for cognition: Developing a six-item version. *Assessment* 27 (8): 1870–85. doi:10.1177/1073191118793208

Degemmis, M., P. Lops, S. Giovanni, M. F. Costabile, L. Oriana, and S. P. Guida. 2004. A hybrid collaborative recommender system based on user profiles. *SCITEPRESS* 5:162–69. doi:10.5220/0002638201620169.

Debczak, M. 2018. New Google Maps feature tells you if you're a 'match' for that restaurant you looked up. July 31. https://www.mentalfloss.com/article/552880/new-google-maps-feature-tells-you-if-youll-actually-restaurant-you-looked.

Epsilon Marketing. 2018. The power of me: The impact of personalization on marketing performan .... Marketing, 15: 51:42 UTC. https://www.slideshare.net/EpsilonMktg/the-power-of-me-the-impact-of-personalization-on-marketing-performance/1.

Eslami, M., S. R. Krishna Kumaran, C. Sandvig, and K. Karahalios. 2018. Communicating algorithmic process in online behavioral advertising. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), Paper no. 432, 1–13. New York, NY: ACM.

Faul, F., E. Erdfelder, A.-G. Lang, and A. Buchner. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39 (2):175–91. doi:10.3758/bf03193146

Hayes, A. F. 2018. Introduction to mediation, moderation, and conditional process analysis. In *A regression-based approach*, 2nd ed. New York: Guilford Publications.

Herlocker, J. L., J. A. Konstan, L. G. Terveen, and J. T. Riedl. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems* 22 (1):5–53. doi:10.1145/963770.963772

Hsiao, K. -L., J. Chuan-Chuan Lin, X. -Y. Wang, H. -P. Lu, and H. Yu. 2010. Antecedents and consequences of trust in online product recommendations: An empirical study in social shopping. *Online Information Review* 34 (6): 935–53. doi:10.1108/14684521011099414

Knijnenburg, B. P., M. C. Willemsen, Z. Gantner, H. Soncu, and C. Newell. 2012. Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction* 22 (4–5):441–504. doi:10.1007/s11257-011-9118-4

Lee, J. D., and K. A. See. 2004. Trust in automation: Designing for appropriate reliance. *Human Factors* 46 (1):50–80. doi:10.1518/hfes.46.1.50_30392

Lee, E.-J., and J. Yoon Jae. 2010. What do others' reactions to news on internet portal sites tell us? Effects of presentation format and readers' need for cognition on reality perception. *Communication Research* 37 (6):825–46. doi:10.1177/0093650210376189

Lynn, M., and H. Judy. 1997. Individual differences in the pursuit of self-uniqueness through consumption. *Journal of Applied Social Psychology* 27 (21):1861–83. doi:10.1111/j.1559-1816.1997.tb01629.x

Netflix. 2020. Netflix ratings & recommendations. Help Center. Accessed April 19, 2020. https://help.netflix.com/en/node/9898.

Ochi, P., S. Rao, L. Takayama and C. Nass. 2010. Predictors of user perceptions of web recommender systems: How the basis for generating experience and search product recommendations affects user responses. *International Journal of Human-Computer Studies* 68 (8):472–82. doi:10.1016/j.ijhcs.2009.10.005

Pearl, P., L. Chen, and H. Rong. 2012. Evaluating recommender systems from the user's perspective: Survey of the state of the art. *User Modeling and User-Adapted Interaction* 22 (4):317–55.

Rentfrow, P. J., L. R. Goldberg, and R. Zilca. 2011. Listening, watching, and reading: The structure and correlates of entertainment preferences. *Journal of Personality* 79 (2): 223–58. doi:10.1111/j.1467-6494.2010.00662.x

See, Y. H. M., R. E. Petty, and L. M. Evans. 2009. The impact of perceived message complexity and need for cognition on information processing and attitudes. *Journal of Research in Personality* 43 (5):880–9. doi:10.1016/j.jrp.2009.04.006

Smith, C. 2019. The future of personalized recommendations is here thanks to the taste app. *KnowTechie*. April 29. https://knowtechie.com/the-future-of-personalized-recommendations-is-here-thanks-to-the-taste-app/.

Snyder, C. R., and H. L. Fromkin. 1977. Abnormality as a positive characteristic: The development and validation of a scale measuring need for uniqueness. *Journal of Abnormal Psychology* 86 (5):518–27. doi:10.1037/0021-843X.86.5.518

Starke, A. D., M. C. Willemsen, and C. C. P. Snijders. 2020. Beyond 'one-size-fits-all' platforms: Applying Campbell's paradigm to test personalized energy advice in the

Netherlands. *Energy Research & Social Science* 59:101311. doi:10.1016/j.erss.2019.101311

Sundar, S. S. 2008. The MAIN model: A heuristic approach to understanding technology effects on credibility. In *Digital media, youth, and credibility*, eds. M. J. Metzger and A. J. Flanagin, 73–100. Cambridge, MA: The MIT Press.

Sundar, S. S., and C. Nass. 2001. Conceptualizing sources in online news. *Journal of Communication* 51 (1):52–72. doi:10.1111/j.1460-2466.2001.tb02872.x

Sundar, S. S., A. Oeldorf-Hirsch, and Q. Xu. 2008. The Bandwagon effect of collaborative filtering technology. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems*, 3453–8. New York: ACM.

Weinschenk, S. M. 2009. *Neuro web design: What makes them click?* 1st ed. Indianapolis, IN: New Riders Publishing.

Zhang, Y., and C. Xu. 2018. Explainable recommendation: A survey and new perspectives. *ArXiv:1804.11192 [Cs]*. https://arxiv.org/abs/1804.11192.

Zillmann, D., and B. Hans-Bernd. 2000. *Exemplification in communication: The influence of case reports on the perception of issues*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.