Author(s): Chen, Sihua; Qiu, Han; Zhao, Shifei; Han, Yuyu; He, Wei; Siponen, Mikko; Mou, Jian; Xiao, Hua

Title: When more is less : the other side of artificial intelligence recommendation

Year: 2022

Version: Published version

Please cite the original version:

Chen, S., Qiu, H., Zhao, S., Han, Y., He, W., Siponen, M., Mou, J., & Xiao, H. (2022). When more is less : the other side of artificial intelligence recommendation. Journal of Management Science and Engineering, 7(2), 213-232. https://doi.org/10.1016/j.jmse.2021.08.001

Journal of Management Science and Engineering 7 (2022) 213–232

Contents lists available at ScienceDirect

Journal of Management Science and Engineering

journal homepage: www.keaipublishing.com/en/journals/journal-of-management-science-and-engineering/

When more is less: The other side of artificial intelligence recommendation

Sihua Chen [a], Han Qiu [a], Shifei Zhao [a], Yuyu Han [a], Wei He [b,*], Mikko Siponen [c], Jian Mou [d], Hua Xiao [e]

[a] School of Information Management, Jiangxi University of Finance and Economics, Nanchang, 330032, China
[b] School of Business Administration, Jiangxi University of Finance and Economics, Nanchang, 330032, China
[c] Faculty of Information Technology, University of Jyvaskyla, Jyvaskyla, 40014, Finland
[d] School of Business, Pusan National University, Busan, 46241, Republic of Korea
[e] Library, Jiangxi Agricultural University, Nanchang, 330045, China

ARTICLE INFO

Article history:
Received 29 January 2021
Received in revised form 12 July 2021
Accepted 1 August 2021
Available online 10 August 2021

Keywords:
AI recommendation
Consumers' preferences
Information cocoon
Consumer decision quality

ABSTRACT

Based on consumers' preferences, AI (artificial intelligence) recommendation automatically filters information, which provokes scholars' debate. Supporters believe that by analyzing the consumers' preferences, AI recommendation enables consumers to choose products more quickly and at a lower cost. Critics deem that consumers are more easily trapped in information cocoons because of the use of AI recommendations. This reduces the possibility of consumers contacting a variety of commodities, thus lowering the consumer decision quality. Based on experiments, this paper discusses the moderating role of AI recommendation on the relationship between consumers' preferences and information cocoons. Moreover, it examines the relationship between information cocoons and consumer decision quality. The findings are: AI recommendation strengthens consumers' preferences; consumers' preferences are positively correlated with information cocoons and further lead to the decline of consumers' decision quality. In the AI era, this paper contributes to revealing the dark sides of AI recommendation and provides empirical evidence for the regulation of AI behaviors.

1. Introduction

Whether the use of AI recommendation technology can help consumers to improve their decision quality is a focus of attention of many scholars (Aksoy et al., 2011). Consumer's decision quality is closely related to their benefits. Meanwhile, it is an important benchmark to evaluate the service quality of electronic markets (Zeithaml et al., 2006). In today's online purchasing platform, consumers are passively using AI recommendation technology. By analyzing the purchasing data, AI recommendation can find consumers' interests and preferences. Based on it, AI recommendation can push the information to consumers and offer product suggestions (Xiao and Benbasat 2007). AI recommendation also contributes to the establishment of online brand loyalty (Xiao and Benbasat 2018). In light of these, the number of online sellers using AI

* Corresponding author.
E-mail address: hewei@jxufe.edu.cn (W. He).

https://doi.org/10.1016/j.jmse.2021.08.001

recommendation technology is surging (Safran and Che 2017). Although AI recommendation has a positive effect on online sellers and contributes to the improvement of purchasing experiences, it triggers people's worries about the information cocoon (Zuiderveen et al., 2016). Consumers' needs for information are in line with consumers' preferences. They tend to selectively get access to information (Xu et al., 2020). AI recommendation processes and provides the purchasing information to the consumers. Moreover, the AI recommendation algorithms incorporate commercial strategies. These commercial strategies also influence the content and order of information the consumers will see (Bozdag 2013; Foster 2012; Schulz et al., 2012; Webster, 2010). Under AI recommendation, the information consumers get access to has already been filtered. Such a filtering algorithm will build an information cocoon. The information that is not in line with the consumer's preferences will be excluded (Tian et al., 2019). Gradually, consumers' preferences will be continuously strengthened. Ultimately, the consumers will bind themselves up in the "information cocoon" (Zhao, 2017).

AI recommendation strengthens the relationship between consumers' preferences and information cocoons. It is difficult for consumers to obtain and use the information that helps consumers make decisions but does not conform to consumers' preferences. It seems that consumers get a lot of AI recommendation information in line with their preferences. But in fact, this homogeneous information in line with their preferences will greatly occupy the cognitive resources of consumers, making them have no energy to obtain other non-homogeneous information in a certain period and eventually, the consumers will be trapped in the information cocoons. As such, the possible categories of products consumers can choose are actually reduced. Some important information that is beneficial for decision is filtered. This may reduce consumers' decision quality. Therefore, more is less.

Currently, the research on AI recommendation and its relationship with consumers' decision quality is insufficient. In the extant literature, some scholars have already explicated that AI recommendation would deepen the restraint degree of information cocoon. Notwithstanding, few scholars further study how to objectively evaluate the restraint degree of information cocoons. Moreover, a vast amount of research has been devoted to the qualitative discussion of the reasons for the formation of information cocoon (Sunstein 2006), the formation process of information cocoon (Guess et al., 2018), and the features of information cocoon (Xu et al., 2020). Based on the qualitative studies of information cocoons, this paper attempts to measure the extent of constriction of information cocoons by combining information entropy and Gini coefficient. Besides, it adopts Delphi method to evaluate the consumers' decision quality.

Based on the above analysis of the relationship of AI recommendation, consumers' preferences, information cocoons and consumer decision quality, this paper attempts to address the following two key problems:

Question 1: Is AI recommendation moderating the relationship between consumers' preferences and information cocoons?
Question 2: Does the information cocoon impact the consumers' decision quality?

To answer these questions, this paper is divided into five sections. The first section is the introduction part, which introduces the background, the insufficiency of literature, the key questions addressed and the innovations. The second section is the theoretical background, which expounds on the model encompassing the constructs of consumers' preferences, AI recommendation, information cocoons, and consumers' decision quality. The third section examines the moderating effect of AI recommendation on the relationship between consumers' preferences and information cocoons. The fourth section tests the relationship between information cocoons and consumers' decision quality. The fifth section is conclusions and outlook.

The innovations of this paper are as follows: first of all, this paper applies the research of information cocoons to consumer online shopping for the first time, trying to provide a different perspective for the study of online consumer decision quality. Second, this paper quantitatively studies the extent of constriction of information cocoons, which provides a new direction for quantifying information cocoons.

2. Literature review and research hypotheses

2.1. Literature review

Information recommendation technology is a kind of method used by the marketing people to enhance the online service. It can effectively improve the information providence efficiency of online sellers (Rust, 2001). Since Goldberg and his colleagues developed the first information recommendation system in 1992, various recommendation systems and related technologies have been introduced to online purchasing platforms. Some researchers believe that AI recommendation technology might produce a negative effect (Benbasat et al., 1990; Todd and Benbasat, 1992; Dabholkar, 2006; West et al., 1999). For instance, AI recommendation may facilitate the formation of information cocoon and constrain consumers in the information cocoon which is in line with their preferences. As such, it deprives consumers of the opportunities to get access to other information (Sunstein, 2006). Notwithstanding, the impact of AI recommendation on the relationship of consumers' preferences and information cocoons still lack empirical proof.

The studies on the relationship between consumers' preferences and information cocoons found that people tried to prove their preexisting preferences by a large amount of internet information. This will strengthen their preferences and accelerate them to form extreme preferences. As a result, it will restrict themselves in the information cocoons (Guess et al., 2018). For example, the program "Daily me" reduces people's access to competitive views and accelerates the polarization of the political attitude of news consumers (Negroponte, 1995). By directly investigating the comprehensive intelligent recommendation systems with various complexities and user modes-the main function of these systems is to provide the products such as films, music and clothes to the consumers, and they drew the following conclusions (Chernev et al., 2015; Johnson et al., 2013; Iyengar and Lepper, 2000):

When consumers are facing too much information, it is daunting for them to find the best products. Therefore, consumers may adopt the potential suboptimal heuristic method, ignore the information that is not in line with their preferences, and hence lead to the formation of information cocoon. Even in the era without recommendation system, the consumers' preferences also prompt them to be blocked into the things that they are interested in and therefore influence their pay intention (Mahajan et al., 1982; Page and Rosenbaum, 1987). Negroponte (1995) held that consumers' preferences lead to the formation of information cocoons. In the study of information cocoon, Zhao found that consumers' preferences were closely related to information cocoons (Zhao, 2017). The continuous strengthening of consumers' references makes the constriction of information cocoons stronger.

The studies on the relationship between information cocoons and consumer decision quality are mainly devoted to the discussion of whether the formation of information cocoons helps the improvement of consumer decision quality. Fleder et al. (2011) considered two opposite directions: fragmentation (that is, an intelligent recommendation system leads to less common points between consumers) and homogenization (that is, intelligent recommendation system leads to the opposite effect of fragmentation). He conducted an empirical study on recommendation systems through group experiments. The conclusions he drew were that: consumers would choose based on their preferences; information cocoons could effectively filter out the information that did not conform to the consumer's preferences; it blocked the consumer's consumption behavior within his preferences. Dellaert (2017) found that the recommendation system promoted consumers to form their own information cocoon, thus helping consumers simplify the decision-making process and reducing the quality of consumers' decision-making at the same time. However, there is still a lack of quantitative research on the impact of information cocoon on the quality of consumer decision-making.

Based on the above literature review, the model explored in this paper is depicted in Fig. 1.

2.2. Research hypotheses

Fleder et al. (2011) believed that AI recommendation facilitates the formation of information cocoons. That is, based on consumers' preferences, AI will recommend relevant information to consumers and facilitate the formation of information cocoons. Moreover, Xu et al. (2020) found that consumers were glad to receive similar information and their information needs were always correlated with their preferences; When consumers conduct online purchase, AI recommendation system will be based on the consumers' purchase information and recommend the information that is in line with consumers' preferences to consumers. If things go on like this, consumers will be trapped in the information cocoons and can't help it. Compared with the era without AI recommendation system, AI recommendation system facilitates the formation process of information cocoons which is based on consumers' preferences. Apart from it, Zhao (2017) found that the constriction of information cocoons would be stronger if the consumers' preferences were continuously strengthened. Hence, this paper proposes the following hypothesis:

H1. AI recommendation positively moderates the relationship between consumers' preferences and information cocoons.

Due to the existence of information cocoons, in the long run, consumers are trapped in their preferences. It is hard for them to obtain opposite or different information (Tian, 2019). When consumers are faced with more purchase information and product categories, information cocoons filter the information that is not in line with consumers' preferences and make consumers hard to make a sound decision. In these circumstances, whether consumers can make an optimal decision depends on how much the optimal decision covers consumers' preferences. Generally, information cocoons reduce the possible samples for consumers to make optimal decisions. Based on the above analysis, this paper proposes the following hypothesis:

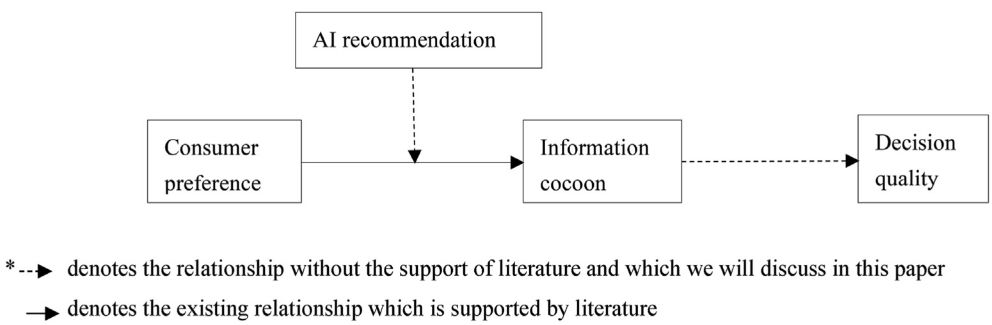

Fig. 1. Framework of the literature review.

H2. information cocoons have a negative impact on consumer decision quality.

Based on the two hypotheses, this paper adopts a multi-group experiment method to simulate the scenario of online purchase. By analyzing the volunteers' historical consumption record, we quantitatively calculate the restraint degree of information cocoons. Then we adopt Delphi method to evaluate the consumer decision quality. Based on it, finally we can statistically analyze the impact of AI recommendation on consumer decision quality.

3. AI recommendation experiment

The purpose of AI recommendation experiment is to test the moderating effect AI recommendation plays on the relationship between consumers' preferences and information cocoons. That is hypothesis 1. The AI recommendation algorithm of each purchase platform is a commercial secret and is hard to obtain. But according to the properties of AI recommendation algorithms and actual testing, we find AI recommendation algorithms have the following characteristics: first, it conducts real-time recommendation based on the data produced by consumers. The AI recommendation ability grows stronger with the increase of consumption data provided by the consumers on the purchase platform. Second, the AI recommendation algorithms of purchase platforms classify the data produced by consumers to the consumers' certain consumption property. Based on millions of consumers' consumption properties, AI can determine a consumer's preferences. To sum up, it is infeasible to decompose the AI recommendation algorithms, extract the consumers' properties we need and analyze each of these property index. To better study the moderating effect of AI recommendation on the relationship between consumers' preferences and information cocoons, we take AI recommendation as a whole to study. In this premise, the experiment is divided into three stages: the pre-experimental stage, the experimental stage and the post experimental stage. In order to verify whether different products and different shopping platforms have the same impact on the experimental results and increase the reliability of the conclusion, two experiments with the same experimental process were organized before and after the experiment: Volunteers in the two experiments purchased different products on two different platforms of Jingdong and Taobao respectively. In the following, the experimental process is described by taking Jingdong platform shopping as an example.

3.1. Pre-experimental stage

The purpose of pre-experimental stage is to make full preparations for the following experiment and to reduce experimental errors. Specifically, this stage consists of determining the product category for experiment, constructing the experimental setting and deciding the subjects of the experiment.

3.1.1. Selection of product

To better examine the moderation effect of AI recommendation, in this experiment we select the products with relatively weak preferences so that the volunteers of the experiment can establish their own preference system. In doing so, the volunteers'preferences are more sensitive to the formation of information cocoons. This will help to better examine the moderation effect of AI recommendation in the relationship between consumers' preferences and information cocoons.

Keyboards are widely used in people's lives. But people's purchase frequency of keyboards is relatively low and a consumer's initial preference toward it is relatively weak. People generally establish their own preferences in the process of purchasing it. So keyboard satisfies the requirement of this experiment. We determine to select keyboard as the product that is designed to purchase. (Similarly, the second experiment selected Taobao platform as the simulation shopping platform and the car tyres were used as the products purchased by the subjects.)

3.1.2. Setting scenario

The experiment needs to simulate the scenario of purchasing keyboard. To improve the external validity, we choose Jingdong, one of the largest electronic equipment online sellers, is the experimental platform. The testees purchase the products on the platform Jingdong and they have no idea whether the computer they used have ever searched the keyboards.

To control the exogenous variation, that is, to test the effect of the only factor of AI recommendation, we choose to implement the experiment in the same course of time, at the same place and in the same environment. Based on the requirements and real situation, the experimental place is set in the computer room of Jiangxi University of Finance and Economics.

3.1.3. Division of groups

We adopt a stratified sampling method to randomly sample the students of different grades at Jiangxi University of Finance and Economics and ensure that the ratio of males to females is about 1:1. To better differentiate the effect of different AI recommendations on the relationship between consumers' preferences and information cocoons, we divide AI recommendations into two types: the ones who have the accounts (that is, the testees log in their accounts to purchase) and the other who have no accounts (that is, the testees' purchase products without accounts). Therefore, we conducted a controlled experiment. We divide the testees into three batches and each batch has about 35 persons.

The first batch of testees simulates purchasing the keyboards on computers without any purchasing data. To ensure the computers these testees use have no purchasing data and recommended information, before the experiment, we formatted and reloaded these computers. AI recommendation is the technology that recommends the information to the consumers based on the items browsed by them. Hence, the influence of AI recommendation on the testees of the first batch is the weakest. This batch is the controlling group and we use $G_{jd}1$ to denote it.

The second batch of testees needs to purchase on the basis of the first batch. That is, the second batch of testees purchase the products on the computers already used by the first batch of testees which have not deleted the browsing information. By now the AI has already made recommendations based on the browsing information of the first batch of testees. Compared with the first batch, the second batch produces more data. Based on the above analysis, the AI recommendation of the second batch is stronger than that of the first batch $G_{jd}1$ and they make contrast with the first batch $G_{jd}1$. So the second batch is the experimental group of $G_{jd}1$ and we use $G_{jd}2$ to denote it. The controlling variable is the amount of purchasing data produced by the testees.

Before the simulation of the purchase of the third batch, we formatted and reloaded the computers. The third batch of testees needs to log in their Jingdong accounts to purchase the products. This design is to test the impact of AI recommendation with an account on the relationship between consumers' preferences and information cocoons. Compared with the first batch, the third batch of testees have logged in their accounts to purchase the products and the AI have already obtained their daily purchasing data. Based on the above analysis, the AI recommendation of the third batch is stronger than that of the first batch $G_{jd}1$ and they make contrast with the first batch $G_{jd}1$. So the third batch is the experimental group of $G_{jd}1$ and we use $G_{jd}3$ to denote it. The controlling variable is whether logging in the Jingdong accounts.

To sum up, $G_{jd}1$ and $G_{jd}2$, $G_{jd}1$ and $G_{jd}3$ respectively form two groups of contrast experiment. Both of them can test the moderating effect of AI recommendation, though the ways to control the degree of AI recommendation are different. In specific, for $G_{jd}1$ and $G_{jd}2$, we control the degree of AI recommendation by controlling the amount of purchasing data produced by the testees. For $G_{jd}1$ and $G_{jd}3$, we control the degree of AI recommendation by controlling whether the testees logged in the Jingdong accounts or not. As such, these two contrast experiments not only test the impact of AI recommendation with different strength degrees on the relationship between consumers' preferences and information cocoons, but also prove the impact of AI recommendation with different categories on the relationship between consumers' preferences and information cocoons.

3.2. The experimental stage

The purpose of the experimental stage is to obtain the real data for each construct and make preparations for the data analysis in the post experimental stage.

3.2.1. Steps of the experiment

(1) conducting the testing of computers and encoding these computers. This experiment uses the same model of computers, encode these computers and test whether these computers are functioning well.

(2) randomly arranging the testees' seats and encoding the testees. When encoding these testees, we make a record of the correspondence between the testees and the computers, that is, for each computer, it corresponds with three testees. This is convenient for the subsequent analysis.

(3) providing guidance for the testees and telling them the procedures of the experiment. To avoid the influence of the guidance on the internal validity, the guidance has not provided information other than the operational process information.

(4) simulating the purchase. Considering that the AI recommendation of Jingdong might be influenced by the computer or browser, to improve the internal validity of the experiment, in this experiment we use the same type of computer devices and the same version of Google browsers. Besides, we simulate these purchases on the same online platform-Jingdong platform. Finally, the testees place the product link and the product screenshot in a specified file, which is used for evaluating the decision quality.

(5) obtaining the testees' historical records of browsing products.

3.2.2. Mathematical modeling analysis

To describe the degree of the strength of information cocoons, we need to quantify the information browsed by the testees while purchasing the products. To quantify the information and describe the total amount of information in a system, Shannon (1948) proposed the concept of information entropy: suppose system X may have n states and they respectively are $x_1, x_2 \ldots x_i \ldots x_n, p(x_i)$ denotes the probability value of the number i state. entropy of the system is defined asThen the information entropy of the system is defined as Eqn (1):

$$H(X) = -\sum_{i=1}^{n} p(x_i) log(p(x_i)) \tag{1}$$

where $0 \leq p(x_i) \leq 1$ and $\sum_{i=1}^{n} p(x_i) = 1$, and we specify when $p(x_i) = 0, 0log0 = 0$.

As a concept of describing the information chaos, the bigger the information entropy, the higher the information chaos. In this experiment, the more types of products browsed by a testee, the higher the information chaos and the bigger the information entropy. According to Sunstein's explanation of information cocoons: in the internet times, people selectively browse and receive the information according to their preferences and interests. As the time goes by, people lose the opportunity to get access to other information and learn about other things. When people are trapped in their preferences and interests, the information cocoons are formed. In the internet times, people who are more constrained by the information cocoons will better classify the information based on its properties. As a result, the information chaos is relatively low and the information entropy is relatively small. To sum up, we believe that information entropy is negatively correlated with the constriction of information cocoons. That is, the bigger the information entropy, the weaker the constriction of the information cocoons.

To overcome the defect of single index in measuring information cocoons, this paper uses Gini impurity to evaluate the extent of information chaos of a system. Similar to information entropy, we suppose that a system X has n states. They are respectively $x_1, x_2 \ldots x_i \ldots x_n$. $p(x_i)$ denotes the probability value of the ith state. Then we define the Gini impurity $G(X)$ of the system as Eqn (2):

$$G(X) = \sum_{i=1}^{n} p(x_i)*(1 - p(x_i)) = 1 - \sum_{i=1}^{n} p(x_i)^2 \tag{2}$$

where $0 \le p(x_i) \le 1$ and $\sum_{i=1}^{n} p(x_i) = 1$. We can know that when the system only has one state, Gini impurity is equal to 0; when there is no big difference in the probability of each state, the more the system states are, the bigger Gini impurity is. Based on the computing formula of Gini impurity, we know that the bigger the Gini impurity, the higher the internal uncertainty of the system. Therefore, Gini impurity is positively correlated with the extent of information chaos in system X. Moreover, the higher the extent of information chaos, the smaller the constriction of the information cocoons. Based on this, we believe that Gini impurity is negatively correlated with the degree of constriction of information cocoons.

In the calculation of information entropy and Gini impurity, p denotes the probability of an event happening. In this experiment, we take all the products a testeen browse as a system. Then the ratio of the time a testee spends on a single product to the total time a testee spend is the probability p. In sum, by using Eqn (1) and (2), we can calculate the degree of information chaos each testee has while browsing the products. This finally reflects the degree of constriction of the information cocoons.

Suppose the total number of items each testee browsed is n. The number m product is k_m. The corresponding browsing time is t_m, then the total amount of time is $T = \sum_{m=i}^{m=n} t_m$. The ratio of the time a testee spends on a single product k_m to the total time a testee spend on all the products is:

$$p(k_m) = t_m / T \tag{3}$$

Combined with Eqn (3), the total information entropy and Gini impurity for the jth testee in the ith batch on the jth computer are:

$$H\left(K_i^{(j)}\right) = -\sum_{m=1}^{n} p(k_m)\log(p(k_m)) \tag{4}$$

$$G\left(K_i^{(j)}\right) = 1 - \sum_{m=1}^{n} p(k_m)^2 \tag{5}$$

There are three batchs of testees. Each batch has N testees. Then $i = 1, 2, 3, j = 1, 2, 3 \ldots N$.

From Eqn (4) and (5) we can know, the information entropy vector and Gini impurity vector of the ith batch of testees are:

$$H(K_i) = \left[H\left(K_i^{(1)}\right), H\left(K_i^{(2)}\right) \ldots H\left(K_i^{(j)}\right) \ldots H\left(K_i^{(N)}\right)\right]_{1XN} \tag{6}$$

$$G(K_i) = \left[G\left(K_i^{(1)}\right), G\left(K_i^{(2)}\right) \ldots G\left(K_i^{(j)}\right) \ldots G\left(K_i^{(N)}\right)\right]_{1XN} \tag{7}$$

From Eqn (6) and (7), we can know the difference vectors of information entropy and Gini impurity of the products browsed by two batchs of testees are Eqn (8) and (9):

$$S(K_i) = H(K_1) - H(K_i) \tag{8}$$

$$D(K_i) = G(K_1) - G(K_i) \tag{9}$$

Eqn (10) and (11) are special situations and we don' t discuss them:

$$S(K_1) = H(K_1) - H(K_1) \tag{10}$$

$$D(K_1) = G(K_1) - G(K_1) \tag{11}$$

Then the internal elements of $S(K_i)$ *and* $D(K_i)$ are Eqn (12) and (13):

$$S\left(K_i^j\right) = H\left(K_1^{(j)}\right) - H\left(K_i^{(j)}\right) \tag{12}$$

$$D\left(K_i^j\right) = G\left(K_1^{(j)}\right) - G\left(K_i^{(j)}\right) \tag{13}$$

If $S(K_i^j) > 0$, then it indicates that the information entropy for the jth testee in the first batch is bigger than the information entropy for the jth testee in the ith batch; if $S(K_i^j) = 0$, then it indicates that the information entropy for the jth testee in the first batch is equal to the information entropy for the jth testee in the ith batch; if $S(K_i^j) < 0$, then it indicates that the information entropy for the jth testee in the first batch is smaller than the information entropy for the jth testee in the ith batch.

Similarly, if $D(K_i^j) > 0$, then it indicates that the Gini impurity for the jth testee in the first batch is bigger than the Gini impurity for the jth testee in the ith batch; if $D(K_i^j) = 0$, then it indicates that the Gini impurity for the jth testee in the first batch is equal to the Gini impurity for the jth testee in the ith batch; if $D(K_i^j) < 0$, then it indicates that the Gini impurity for the jth testee in the first batch is smaller than the Gini impurity for the jth testee in the ith batch.

We define the signum function according to $S(K_i^j)$ *and* $D(K_i^j)$ as Eqn (14) and (15).

$$sig\left(K_i^j\right) = \begin{cases} -1, \ S\left(K_i^j\right) < 0 \\ 1, \ S\left(K_i^j\right) > 0 \\ 0, \ S\left(K_i^j\right) = 0 \end{cases} \tag{14}$$

$$sigd\left(K_i^j\right) = \begin{cases} -1, \ D\left(K_i^j\right) < 0 \\ 1, \ D\left(K_i^j\right) > 0 \\ 0, \ D\left(K_i^j\right) = 0 \end{cases} \tag{15}$$

On the premise that the data produced by the testees are real and valid, to avoid the influence of extreme data on the whole experimental data but not deleting these extreme data, we apply the signum function to compare the information entropies of the first batch of testees and those of the ith batch of testees. That is, we use the following formula to compare these information entropies.

$$LS(K_i) = \sum_{j=1}^{N} sigs\left(K_i^j\right) \tag{16}$$

$$LD(K_i) = \sum_{j=1}^{N} sigd\left(K_i^j\right) \tag{17}$$

From Eqn (16), we can know if $LS(K_i) > 0$, then it indicates that the information entropies of the first batch of testees are bigger than those of the ith batch of testees. The information chaos is relatively high and the constriction of the information cocoons is relatively weak; if $LS(K_i) < 0$, then it indicates that the information entropies of the first batch of testees are smaller than those of the ith batch of testees. The information chaos is relatively low and the constriction of the information cocoons is relatively strong; if $LS(K_i) = 0$, then it indicates that the information entropies of the first batch of testees are equal to those of the ith batch of testees. The information chaos is equal and the constriction of the information cocoons is equal. It is similar to Eqn (17).

3.3. Post experimental stage

The main purpose of the post experiment stage is to establish a mathematical model to analyze the data. By solving the model, we can further test hypothesis 1, that is, whether there is a moderating effect of AI recommendation on the relationship between consumers' preferences and information cocoons.

3.3.1. Data normalization

Before formally processing the experiment data, we preprocess the data and delete the data that doesn't match the requirement. The principles are: (1) for those we can't obtain the historical browsing records due to the testees'operational problems, we delete them; (2) for those who haven't filled in the purchase file, that is, we can't make an evaluation of the decision quality, we delete them; (3) for those who violate the experimental operation procedures and then lead to the involvement of other variables, we delete their data; (4) in the experiment, the testees randomly use the computers. We specify that the testees who use the same computers are of the same group. Then if we delete a testee's data, we delete the data of the whole group where the testee is in.

According to the delete principles, except for those which don't match the principles, finally we altogether have 28 groups and each group has 3 persons in the experiment on Jingdong platform and have 32 groups and each group has 3 persons in the experiment on Taobao platform.

3.3.2. Data process and verification of hypotheses

According to the real and valid data which is already normalized, we adopt Eqn (1) and (2) to calculate the information entropies and Gini impurity. In the Jingdong platform experiment and Taobao platform experiment, the values of information entropy and Gini impurity for each testee please see appendix Table A.1 and Table A.2, where the first digital of ID denotes the ith batch and the last two digitals denote the jth testee. To show the results more clearly, the values of information entropy and Gini impurity are plotted. The point-fold line charts show the values of information entropy and Gini impurity of different batches of testees. The histograms show the sum values of information entropy and sum values of Gini impurity of different batches of testees. The values of information entropy of Jingdong platform please see Fig. 2a.

From Fig. 2a we can see that in the Jingdong platform the information entropies of $G_{jd}1$ are higher than those of $G_{jd}2$ generally; almost all the information entropies of $G_{jd}1$ are higher than those of $G_{jd}3$; that is, $T(G_{jd}1) > T(G_{jd}2)$、$T(G_{jd}1) > T(G_{jd}3)$. In light of that the information entropy is negatively correlated with the information cocoons, that is, the bigger the information entropy, the weaker the constriction of information cocoons. Then we further know that the constriction of information cocoons of $G_{jd}1$ is weaker than that of $G_{jd}2$, $G_{jd}3$. In this experiment, the level of AI recommendation of $G_{jd}2$, $G_{jd}3$ are higher than $G_{jd}1$.

The Gini coefficient statistics chart of Jingdong platform volunteers is shown in Fig. 2b below.

From Fig. 2b we can see that on the Jingdong platform the Gini impurities of $G_{jd}1$ are all higher than those of $G_{jd}2$; almost all the Gini impurities of $G_{jd}1$ are higher than those of $G_{jd}3$; that is, $T(G_{jd}1) > T(G_{jd}2)$、$T(G_{jd}1) > T(G_{jd}3)$. In light of that the Gini impurity is negatively correlated with the information cocoons, that is, the bigger the Gini impurity, the weaker the constriction of information cocoons. Then we further know that the constriction of information cocoons of $G_{jd}1$ is weaker than that of $G_{jd}2$, $G_{jd}3$. In this experiment, the level of AI recommendation of $G_{jd}2$, $G_{jd}3$ are higher than $G_{jd}1$.

Similarly, the information entropy and Gini impurity of each group of testees on the Taobao platform are calculated and plotted. The statistical charts of information entropy and Gini impurity of Taobao platform are shown in Fig. 3a and Fig.3b below.

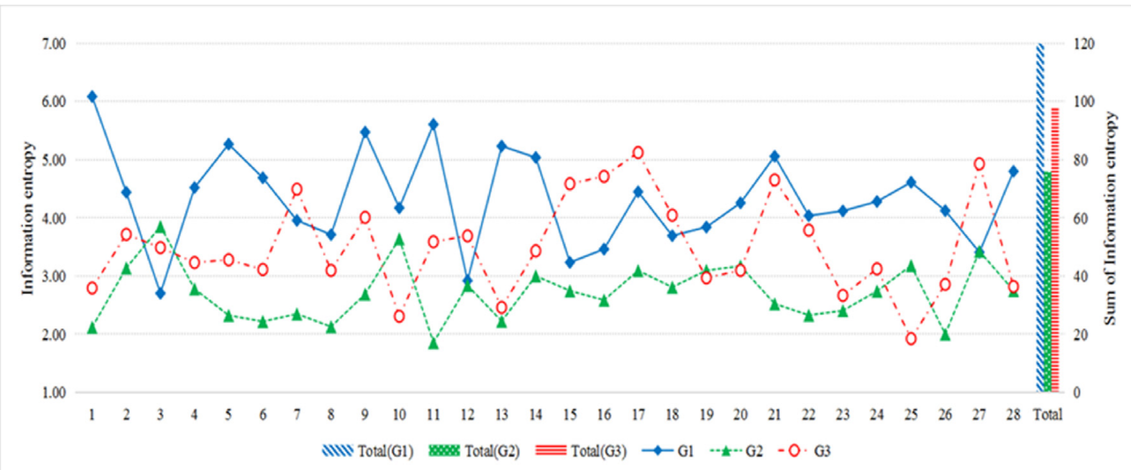

Fig. 2a. point-fold line chart of information entropy of different batches of testees on Jingdong platform.

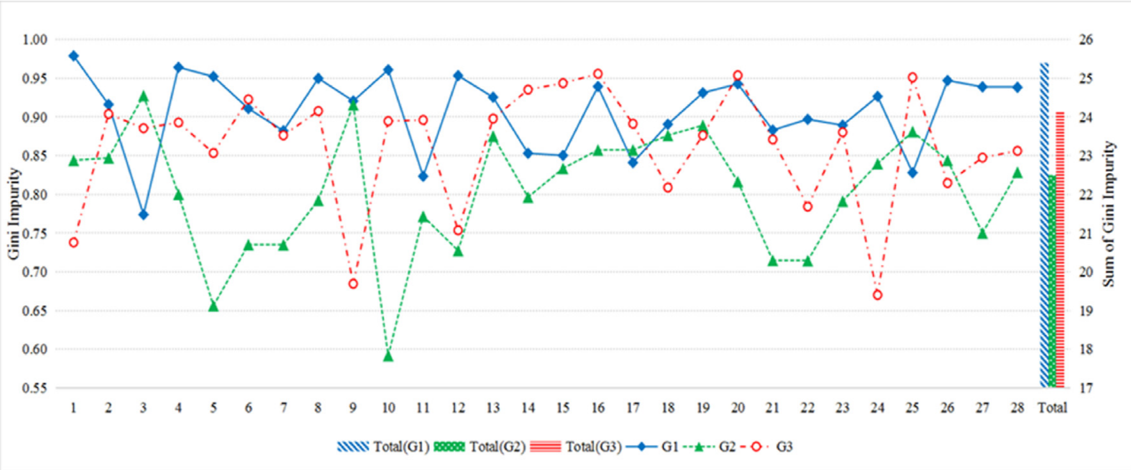

Fig. 2b. point-fold line chart of Gini impurity of different batches of testees on Jingdong platform.

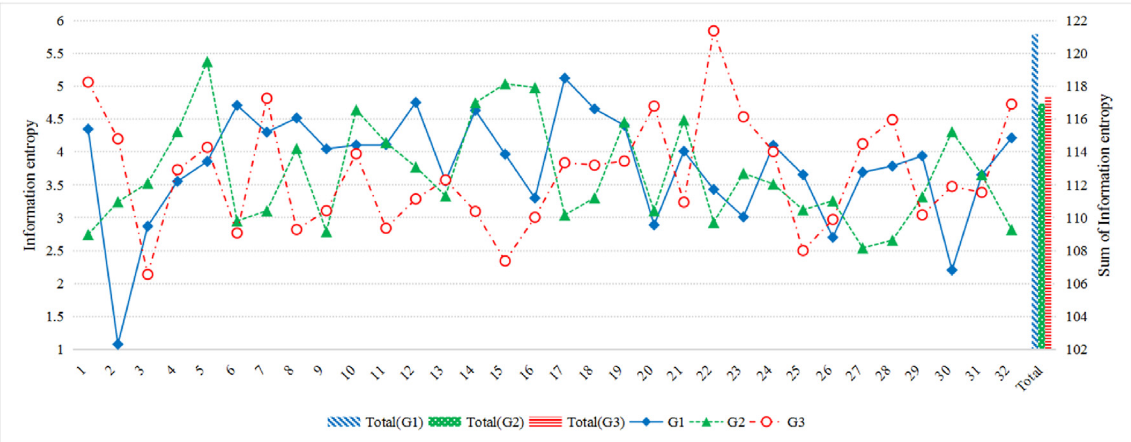

Fig. 3a. point-fold line chart of information entropy of different batches of testees on Taobao platform.

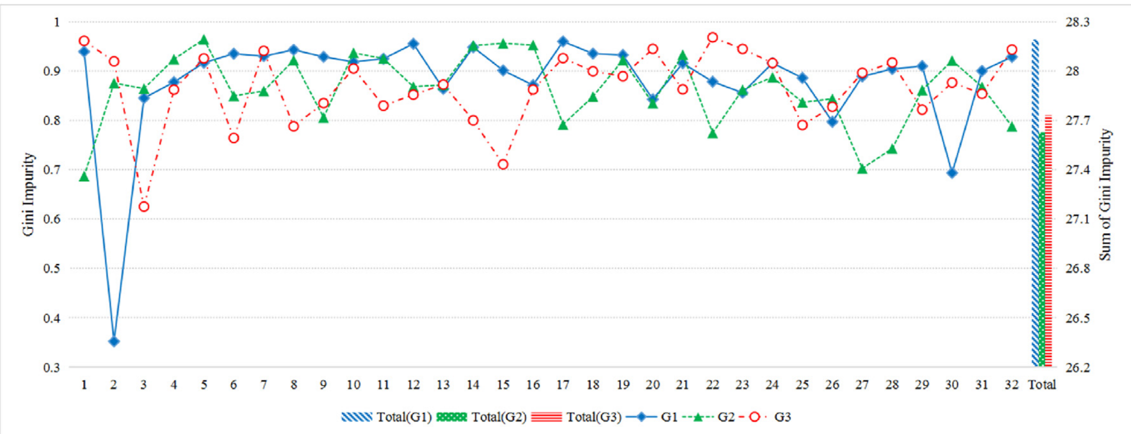

Fig. 3b. point-fold line chart of Gini impurity of different batches of testees on Taobao platform.

Compared with Fig. 2a and b, we can see that the point-fold lines of information entropy and Gini impurity of different batches of testees on Taobao platform intertwined and it is hard to tell the differences visually. As such, we compute the sum of information entropies and sum of Gini impurity and get the results:$T(\mathrm{G_{tb}}1) > T(\mathrm{G_{tb}}2)$、$T(\mathrm{G_{tb}}1) > T(\mathrm{G_{tb}}3)$. Therefore, we know that the constriction of information cocoons of $\mathrm{G_{tb}}1$ is weaker than that of $\mathrm{G_{tb}}2$, $\mathrm{G_{tb}}3$. The level of AI recommendation of $\mathrm{G_{tb}}2$, $\mathrm{G_{tb}}3$ are higher than $\mathrm{G_{tb}}1$.

Based on the information entropy and Gini impurity of Jingdong platform and Taobao platform, we can make a preliminary verification of hypothesis 1: AI recommendation plays a moderating role in the relationship between consumers' preferences and information cocoons. In other words, under a certain degree of consumers' preferences, the higher the degree of AI recommendation, the stronger the constriction of information cocoons.

3.3.3. Further verification of hypotheses

To prevent the influence of some extreme data, we use Eqn (16) and (17) to restrict the extreme data in its group. Based on above equations, we calculate $\mathrm{LS_{jd}}(K_i)$ and $\mathrm{LD_{jd}}(K_i)$ and please see Table 1.

From Table 1 we know that the value of $\mathrm{LS_{jd}}(K_2)$ is 24. It indicates that in 28 groups of testees, the information entropies of 26 testees of $\mathrm{G_{jd}}1$ exceed the information entropies of 26 testees of $\mathrm{G_{jd}}2$. The value of $\mathrm{LS_{jd}}(K_3)$ is 12. It indicates that in 28 groups of testees, the information entropies of 20 testees of $\mathrm{G_{jd}}1$ exceed the information entropies of 20 testees of $\mathrm{G_{jd}}3$. The value of $\mathrm{LD_{jd}}(K_2)$ is 22. It indicates that in 28 groups of testees, the Gini impurity of 25 testees of $\mathrm{G_{jd}}1$ exceeds the Gini impurity of 25 testees of $\mathrm{G_{jd}}2$. The value of $\mathrm{LD_{jd}}(K_3)$ is 10. It indicates that in 28 groups of testees, the Gini impurity of 19 testees of $\mathrm{G_{jd}}1$ exceeds the Gini impurity of 19 testees of $\mathrm{G_{jd}}3$. The values of $\mathrm{LS_{jd}}(K_i)$ and $\mathrm{LD_{jd}}(K_i)$ are both bigger than 0. Based on the above analysis, we know that the information entropies and Gini impurity of $\mathrm{G_{jd}}1$ are greatly bigger than those of $\mathrm{G_{jd}}2$ and $\mathrm{G_{jd}}3$. It indicates that the degree of information chaos is higher and the constriction of information cocoons is weaker.

Similarly, we use Eqn (16) and (17) to restrict the extreme data in its group. Based on above equations, we calculate $\mathrm{LS_{tb}}(K_i)$ and $\mathrm{LD_{tb}}(K_i)$ and please see Table 2.

From Table 2 we know that the value of $\mathrm{LS_{tb}}(K_2)$ is 0. It indicates that in 32 groups of testees, the information entropies of $\mathrm{G_{tb}}1$ is equal to those of $\mathrm{G_{tb}}2$. The value of $\mathrm{LS_{tb}}(K_3)$ is 4. It indicates that in 32 groups of testees, the information entropies of 18 testees of $\mathrm{G_{tb}}1$ exceed the information entropies of 18 testees of $\mathrm{G_{tb}}3$. The value of $\mathrm{LD_{tb}}(K_2)$ is 4. It indicates that in 32 groups of testees, the Gini impurity of 18 testees of $\mathrm{G_{tb}}1$ exceed the Gini impurity of 18 testees of $\mathrm{G_{tb}}2$. The value of $\mathrm{LD_{tb}}(K_3)$ is 4. It indicates that in 32 groups of testees, the Gini impurity of 18 testees of $\mathrm{G_{tb}}1$ exceed the Gini impurity of 18 testees of $\mathrm{G_{tb}}3$. The values of $\mathrm{LS_{tb}}(K_i)$ and $\mathrm{LD_{tb}}(K_i)$ are both bigger or equal to 0. Based on the above analysis, we know that the information entropies and Gini impurity of $\mathrm{G_{tb}}1$ are generally bigger than those of $\mathrm{G_{tb}}2$ and $\mathrm{G_{tb}}3$. It indicates that the degree of information chaos is higher and the constriction of information cocoons is weaker.

From the above analysis, we can see that, on the whole, on the platforms of Jingdong and Taobao, the information the testees browse after the information recommendation is more concentrated.

Based on the above analysis, the experiment further verifies hypothesis 1: AI recommendation moderates the relationship of consumers' preferences and information cocoons. That is, given consumers' preferences, the higher AI recommendation, the stronger the constriction of information cocoons. The results are consistent with those drawn from subsection 3.3.2.

4. Analysis of decision quality based on Delphi method

Delphi method raises questions to experts anonymously and conducts three or four rounds of consultation. After repeated anonymous consultations, the opinions tend to be consistent and the final results are obtained. As such, Delphi method is suitable for analyzing the problems with large fuzziness and the problems that are unable to make quantitative analysis directly. In this paper, it is hard to quantify the consumers decision quality. Because of the characteristics analyzed above, we adopt Delphi method to evaluate the consumer decision quality. Based on experts' repeated analysis and feedback, we establish an index evaluation system of keyboard quality. By scoring the purchased keyboards, we compared the decision quality of the three batches of testees.

Table 1
Statistical table for $\mathrm{LS_{jd}}(K_i)$、$\mathrm{LD_{jd}}(K_i)$ of Jingdong platform.

ID	Value	ID	Value
$\mathrm{LS_{jd}}(K_2)$	24	$\mathrm{LD_{jd}}(K_2)$	22
$\mathrm{LS_{jd}}(K_3)$	12	$\mathrm{LD_{jd}}(K_3)$	10

Table 2
Statistical table for $\mathrm{LS_{tb}}(K_i)$ *and* $\mathrm{LD_{tb}}(K_i)$ of Taobao platform.

ID	Value	ID	Value
$\mathrm{LS_{tb}}(K_2)$	0	$\mathrm{LD_{tb}}(K_2)$	4
$\mathrm{LS_{tb}}(K_3)$	4	$\mathrm{LD_{tb}}(K_3)$	4

4.1. Identifying experts

The selection of experts is the first step of Delphi method. It is directly related to the accuracy of the evaluation. To establish the quality evaluation system of "keyboard" and "tyre" products, we select the experts mainly based on industries, working years, frequencies and categories of keyboard and tyre purchase, and keyboard uses. Finally, a total of 9 experts were selected and the detailed information of them is shown in Table 3.

4.2. Establishing the quality evaluation system for "keyboard" and "tyre" products

This experiment conducted three rounds of Delphi expert consultation by filling in the questionnaire. The first round of expert consultation initially drew up the evaluation index of keyboard and tyre quality through the expert's reply; the second round of expert consultation made the experts evaluate the importance of the preliminary evaluation index and put forward the modification opinions of the index; the third round of expert consultation made the experts evaluate the index again by comprehensively referring to the opinions of others. After the third round of evaluation forms was withdrawn, the experts' opinions tended to be consistent and reliable so as to establish the quality evaluation index system of "keyboard" and "tyre" products. In this experiment, 9 copies of consultation forms were sent out in the first round, and 9 copies were recovered, with the recovery rate of 100%; in the second round, 9 copies were sent out and 9 copies were recovered, with the recovery rate of 100%; in the third round, 9 copies were sent out, and the recovery rate was 100%. The specific flow chart is shown in Fig. 4.

4.2.1. The first round of consultation

After the experts were selected, we launched the first round of e-mail consultation for the experts. In the first round of consultation, the experts were informed of the background and purposes of the experiment. We constructed a purchase scenario for the experts and required the experts to list the indicators to judge whether the keyboard was good or bad when they bought the keyboard. After collecting the answers, we summarized the experts' answers. Similarly, the summary of tire expert opinion is shown in Table 4.

By summarizing the opinions of keyboard experts, three one-level indicators and 14 two-level indicators are drawn up, and the keyboard quality evaluation indexes are preliminarily established. In the same way, four one-level indicators and 17 two-level indicators are drawn up, and the tyre quality evaluation indexes are established preliminarily. The specific indicators are shown in Table 5.

4.2.2. The second round of consultation

After the first round of expert consultation, the preliminary evaluation index of keyboard quality is established. Then the indicators were sent to experts and the second round of expert consultation was launched. Experts evaluated the importance of the preliminary evaluation index according to a Likert's 5-point system. Among them, 1 is very unimportant, 2 is unimportant, 3 is indifferent, 4 is important and 5 is very important. Meanwhile, experts were required to put forward suggestions on the revision of the index. The experts' opinions were collected. Four of them questioned the indicator of "top search ranking"; three of them questioned the indicator of "waterproof and dustproof" and "anti-skid material"; and three questioned the indicators of "preferential activities" and "unique design". The experts believed that these indicators should not be used to evaluate the quality of keyboards.

Similarly, the opinions of tyre experts were summarized. Four of them questioned the indicator of "beautiful", two questioned the indicator of "after-sale services", and one questioned the indicators of "brand", "high temperature resistance" and "speed limit".

4.2.3. The third round of consultation

After the second round of expert consultation, the opinions were summarized and sent to each expert anonymously so that the experts could refer to the opinions of other experts to score the indicators again. Finally, the expert opinions tend to be consistent and reliable. The indexes are finally determined by combining the statistical index and the experts' opinions. Next, the calculation of statistical indicators is described.

(1) Mean and frequency of perfect score

To get the concentration degree of expert opinions, we use the following Eqn (18) and (19) to calculate the indicators of mean and frequency of full mark of the data.

$$M_a = \frac{1}{n_a}\left(\sum_{z=1}^{n} C_{za}\right) \tag{18}$$

Where n_a denotes the number of experts who participated in the evaluation of the ath indicator; C_{za} denotes the evaluation score of the zth expert on the ath indicator. The bigger M_a is, the more important ath indicator is.

Table 3
Information of experts identified for the product purchase.

A Information of experts identified for keyboard purchase				
Category	Industry	Working year	Frequency of keyboard purchase	number
Keyboard seller	Keyboard salesman	More than 15 years	More than 50 times	2 persons
Professional long term users	Programmer, software engineer, software operation and maintenance	More than 10 years	More than 10 times	4 persons
E-sports fans	PE investment, employees in financial companies, civil servants	More than 10 years	More than 10 times	3 persons
B Information of experts identified for tyre purchase				
Category	Industry	Working year	Frequency of tyre purchase	number
Tyre seller	Auto after-salesman, auto salesman	4–12 years	2-5 times	6 persons
Engineer	Auto parts after-sales engineer, parts R&D engineer, 4S shop repairman	4–18 years	1-5 times	3 persons

$$S_a = \frac{n'_a}{n_a} \tag{19}$$

Where n_a denotes the number of experts who participated in the evaluation of the ath indicator; n'_a denotes the number of experts who give full scores. The values of S_a between 0 and 1 and S_a can be used as the supplementary indicator of M_a. The bigger S_a, which indicates more experts giving full scores to the indicator, then the greater importance the indicator.

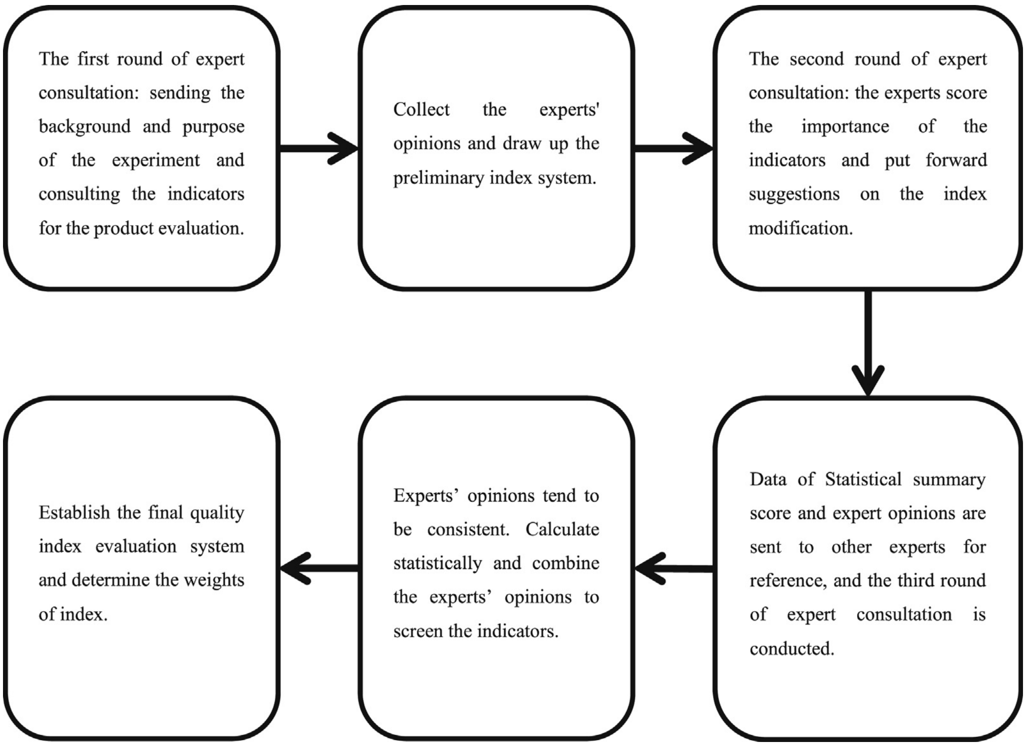

Fig. 4. Flow chart of establishment of the product quality evaluation index system.

Table 4
Summary of expert index.

A Summary of keyboard expert index	
Expert number	**Indicator**
Expert 1	Cool, easy to press, sensitive, convenient, foldable, cheap, waterproof, dustproof, lightweight
Expert 2	Easy to operate, comfortable, sensitive, beautiful design, good material, function key, long service life, easy to clean, dustproof, waterproof, wear-resistant
Expert 3	Appearance, sensitivity, feeling of touch, price, practicability, after-sale service, discount and reputation
Expert 4	Appearance, feeling of touch, material, performance, quality and value for money, brand awareness, after-sales service, attitude of salesmen
Expert 5	Shape design, price, size, portability, sense of operation, material, service life and positioning
Expert 6	Brand, price, preferential activities, evaluation, online customer service response timely, top search ranking, favourable rate
Expert 7	Brand, quality, price, feeling of touch, appearance, dustproof, splash proof, antiskid materials, etc
Expert 8	Feeling of touch, appearance, size, power consumption
Expert 9	Feeling of touch, beautiful appearance, brand, price, function, wireless, durable, color
B Summary of tyre expert index	
Expert number	**Indicator**
Expert 1	Safety, brand, maintenance, wear resistance, price, convenience of purchase, on-site installation and convenience of customer service
Expert 2	Wear resistance, anti-slip, noise, brand, price, after-sales, tread, side extrusion
Expert 3	Safety, price, comfort, brand, wear resistance, tire noise, load bearing, speed limit
Expert 4	Grip, wear resistance, silence, shock absorption, handling, high temperature resistance, price, rolling resistance, wetland grip
Expert 5	Noise, wear resistance, smoothness, brand, price, after-sale services
Expert 6	Safety, comfort, quiet, wear resistance, economic, good water conductivity, beautiful
Expert 7	Tire explosion proof, wear resistance, price, grip, flat ratio, brand, matching with car
Expert 8	Wear resistance, safety, manipulation, price, energy saving, service, noise, brand
Expert 9	Size, grip, wear resistance, antiskid, antifreeze, sunscreen, price, aging

Table 5
Preliminary indicators of evaluation of product quality.

A Preliminary indicators of evaluation of keyboard quality	
One-level indicator	**Two-level indicator**
quality	Material stand wear and tear
	Anti-skid materials
	Waterproof and dustproof
	Service life
	Beautiful appearance
	Unique design
practicability	Feel comfortable
	The key is sensitive
	Quality and value for money
	Portable and foldable
brand	Customer service
	Preferential activities
	Top search ranking
	Favourable rate
B Preliminary indicators of evaluation of tyre quality	
One-level indicator	**Two-level indicator**
Safety	Water conductivity
	Load bearing
	Speed limit
	Grip
	Manipulation
	Tyre explosion proof
Quality	Wear resistance
	High temperature resistance
	Antifreeze
Comfort	Low tyre noise
	Shock absorption
	Energy saving
	Smoothness
Other factors	Brand
	Customer service/after-sale services
	Performance/cost ratio
	Beautiful

(2) Variation coefficient and coordination coefficient

We calculate the variation coefficient and coordination coefficient to evaluate the coordination of expert opinions. Based on these coefficients, we can judge whether there are big differences in the evaluation of each indicator by experts.

The variation coefficient is Eqn (20):

$$V_a = \frac{\delta_a}{M_a} \tag{20}$$

Where V_a denotes the variation coefficient of the ath indicator; δ_a denotes the standard deviation of the ath indicator. The variation coefficient indicates the evaluation fluctuation of the relative importance of the ath indicator or the coordination. That is, the smaller V_a, the higher the coordination of experts' opinions.

The coordination coefficient(W) indicates the degree of consistency of expert opinions on each indicator and whether there are big differences in opinions on each indicator. It is the credibility indicator of the consultation results. The coordination coefficient indicates the coordination of n experts on b indicators. The values of coordination coefficient are between 0 and 1. The bigger the coordination coefficient, the higher the coordination of expert opinions. Vice versa. Generally speaking, after two to three rounds of consultations, the coordination coefficient will be bigger than 0.5 and approach 1. The coordination coefficient is Eqn (21):

$$W = \frac{12}{n^2\left(b^3 - b\right) - n\sum_{z=1}^{n} T_z} \sum_{a=1}^{b} d_a^2 \tag{21}$$

Where n denotes the number of experts;b denotes the number of indicators;d_a denotes the difference in the grade of indicator a and the arithmetic average value of the sum of all the index grades; T_z denotes the modified coefficient.

From Table 6 we can see that through the second round and third round of consultation, experts' understanding of the importance of indicators tends to be consistent and reliable. The p-values of chi square test of two rounds of coordination coefficient were all less than 0.05, which indicated that under the 95% confidence level, the expert evaluation opinions were well coordinated and the results were acceptable.

(3) Screening evaluation index

Based on the calculation of the statistical indicators such as mean, frequency of full mark, standard deviation, and variation coefficient, we ranked the relative importance of the indicators. Based on the experts' opinions in the consultation process, we determined the final product quality evaluation index system. We also calculated the weight of each indicator. The relative importance of each index is shown in Table 7.

According to the ranking of each indicator and the experts' opinions, we screened out six indexes, which are the top search ranking, anti-skid materials, unique design, waterproof and dust-proof, portable and foldable and preferential activities. Eventually we obtained the final selected indicators. So far, we have established a keyboard quality evaluation system.

According to the ranking results of various indicators of tyre and the expert opinions in the consultation process, we screened out three indicators of beautiful, brand and customer services/after-sale services, and got the final selected indicators. For other factors, there is only one two-level index “performance/cost ratio”. Hence, we classified this indicator into quality category. So far, we have established a tyre quality evaluation system.

Table 6
Statistical tables for coordination coefficient of experts.

	A Statistical table for coordination coefficient of keyboard experts	
	Second round of consultation	**Third round of consultation**
Number of indicators	9	9
Coordination coefficient	0.24	0.726
Chi square value	27.70	84.905
p-value	0.01	0.00
	B Statistical table for coordination coefficient of tyre experts	
	Second round of consultation	**Third round of consultation**
Number of indicators	9	9
Coordination coefficient	0.286	0.500
Chi square value	41.15	72.021
p-value	0.01	0.00

Table 7
The relative importance of each indicator for the products.

Indicator	R_{M_a}	R_{δ_a}	R_{S_a}	R_{V_a}	Expectation	Ranking
A The relative importance of each indicator for keyboard						
Material stand wear and tear	1	1	1	1	1	1
The key is sensitive	2	2	2	2	2	2
Feel comfortable	3	3	3	3	3	3
Service life	5	3	5	4	4.25	4
Favourable rate	5	3	5	4	4.25	4
Quality and value for money	4	8	4	6	5.5	6
Customer service	7	3	8	7	6.25	7
Beautiful appearance	7	3	8	7	6.25	7
top search ranking	7	12	7	10	9	9
anti-skid materials	10	10	8	9	9.25	10
unique design	13	9	8	11	10.25	11
waterproof and dust-proof	12	11	8	12	10.75	12
portable and foldable	11	13	8	13	11.25	13
preferential activities	14	14	8	14	12.5	14
B The relative importance of each indicator for tyre						
Grip	1	2	1	2	1.5	1
Wear resistance	1	2	1	2	1.5	1
Performance/cost ratio	4	2	6	4	4	3
Manipulation	4	6	4	5	4.75	4
Water conductivity	6	1	12	1	5	5
Tyre explosion proof	3	15	1	11	7.5	6
Load bearing	10	6	10	6	8	7
Energy saving	12	9	10	7	9.5	8
Shock absorption	6	14	6	13	9.75	9
Low tyre noise	12	9	12	7	10	10
Smoothness	14	6	12	9	10.25	11
Speed limit	10	12	8	12	10.5	12
High temperature resistance	6	17	4	16	10.75	13
Antifreeze	9	13	8	14	11	14
Beautiful	17	5	12	10	11	14
Brand	16	11	12	15	13.5	16
Customer services/after-sale services	15	16	12	17	15	17

4.2.4. Determination of weights

Based on the proportion of the expert's score on the index, the weight of each index is calculated. On the basis of it, we can calculate the evaluation score of the products the testees purchased in the subsequent experiment. The weight of each index is calculated as Eqn (22):

$$\lambda_q = \frac{B_q}{\sum B_q} \tag{22}$$

where q denotes the first-level indicator; λ_q denotes the weight of the qth indicator; B_q denotes the sum of expert score of the qth first-level indicator; B_a denotes the average score of experts of the ath second-level indicator; $B_q = \sum B_{qa}; \sum B_q$ denotes the sum of the average scores of experts for all indicators. According to the above equations, we calculate the weights of indicators. Please see Table 8.

4.3. Scoring of keyboard decision quality

After the establishment of the product quality evaluation index system and giving weight to each indicator, we carry out the final step of the experiment: to score the product. The experts scored the three batches of products that the subjects finally decided. The score of each product was composed of the scores of the eight indicators established above. The scoring of each indicator was conducted by a Likert five score system, in which 1 was very dissatisfied, 2 was dissatisfied, 3 was average, 4 was satisfactory, and 5 was very satisfied. Then, we calculate the final score of each product according to the weight of each indicator. The specific equations are (23) and (24):

$$Y_{ert} = \sum \lambda_q g_{ertq} \tag{23}$$

$$Y_{rt} = \frac{\sum_{e=1}^{9} Y_{ert}}{9} \tag{24}$$

Where e denotes the experts, r denotes the batch of products, t denotes the number of each batch; g_{ertq} denotes the average score of expert e for the qth first-level indicator of the rth batch of product with the tth number;Y_{ert} denotes the final score of

Table 8
Indicator weights for the products.

One-level indicator	Two-level indicator	Value	Weight
	A Indicator weights for keyword product		
quality	Material stand wear and tear	4.67	37%
	Service life	4.33	
	Beautiful appearance	4.22	
practicability	Feel comfortable	4.44	38%
	The key is sensitive	4.67	
	Quality and value for money	4.44	
brand	Customer service	4.33	25%
	Favourable rate	4.56	
	B Indicator weights for tyre product		
One-level indicator	**Two-level indicator**	**Value**	**Weight**
Safety	Water conductivity	4.00	44.42%
	Load bearing	3.67	
	Speed limit	3.67	
	Grip	4.67	
	Manipulation	4.33	
	Tyre explosion proof	4.44	
Quality	Wear resistance	4.67	30.08%
	High temperature resistance	4.00	
	Antifreeze	3.78	
	Performance/cost ratio	4.33	
Comfort	Low tyre noise	3.44	25.50%
	Shock absorption	4.00	
	Energy saving	3.44	
	Smoothness	3.33	

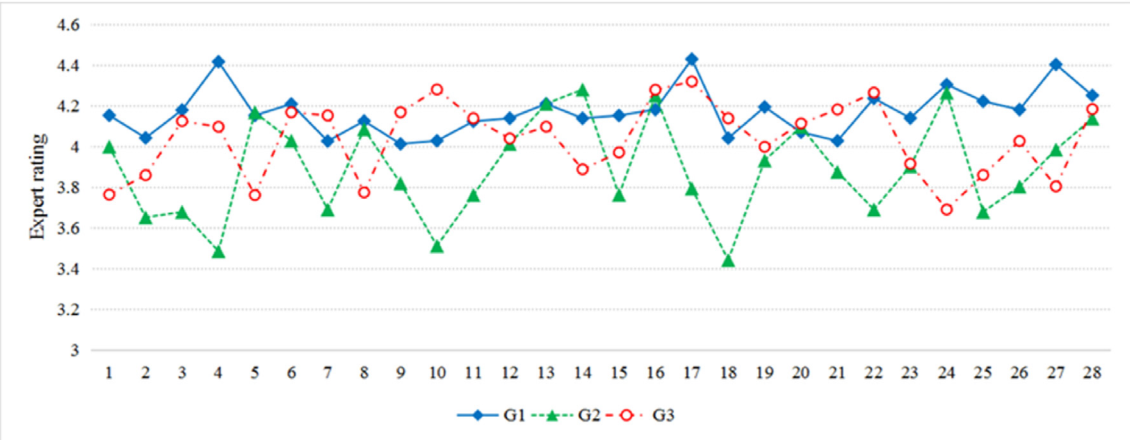

Fig. 5a. Point-fold line chart of keyboard expert scoring of different batches.

expert e for the rth batch of product with the tth number; Y_{rt} denotes the final average score of all experts for the rth batch of product with the tth number. By calculating the score of the nine experts, the final scores for the products purchased by testees are shown in Fig. 5a and b.

4.4. Comparison of decision quality

After calculating the value for each decision product, we compare the quality of these products. Based on Eqn (6), we can get the following Eqn (25):

$$C\left(K_i^j\right) = D\left(K_1^{(j)}\right) - D\left(K_i^{(j)}\right) \tag{25}$$

where $C(K_i^j)$ denotes the difference between the first batch of testees and the ith batch of testees on the jth product quality. If $C(K_i^j) > 0$, then it indicates the score for the jth product of the first batch is bigger than the score for the jth product of the ith batch; if $C(K_i^j) = 0$, then it indicates the score for the jth product of the first batch is equal to the score for the jth product of the ith batch; if $C(K_i^j) < 0$, then it indicates the score for the jth product of the first batch is smaller than the score for the jth product of the ith batch.

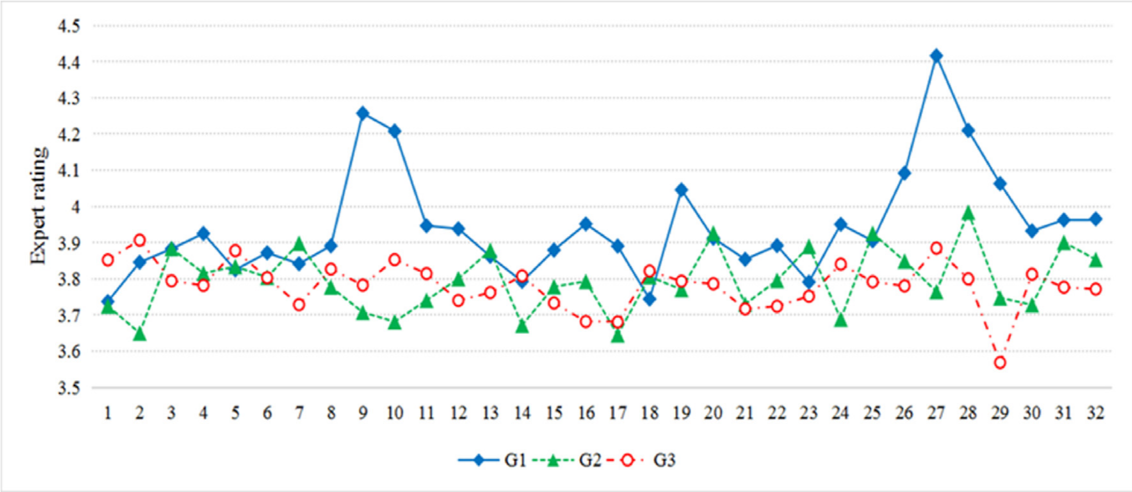

Fig. 5b. Point-fold line chart of tyre expert scoring of different batches.

In order to avoid the effect of extreme data on the product quality, we use the following Eqn (26) and (27) to compare the product quality of the first batch and that of the ith batch.

$$sig\left(K_i^j\right) = \begin{cases} 1, \ C\left(K_i^j\right) > 0 \\ 0, \ C\left(K_i^j\right) = 0 \\ -1, \ C\left(K_i^j\right) < 0 \end{cases} \tag{26}$$

$$E(K_i) = \sum_{j=1}^{N} sig\left(K_i^j\right) \tag{27}$$

If $E(K_i) > 0$, it indicates that the decision quality of the first batch of testees is generally better than that of the ith batch of testees; if $E(K_i) = 0$, it indicates that the decision quality of the first batch of testees is generally equal to that of the ith batch of testees; if $E(K_i) < 0$, it indicates that the decision quality of the first batch of testees is generally worse than that of the ith batch of testees. The final results are shown in Table 9.

As can be seen from the data in Table 9, the value of $E_{\mathrm{jd}}(K_2)$ of keyboard was 18, which indicated that among 28 groups of experts, 23 groups of the first batch had higher scores than those of the second batch; the value of $E_{\mathrm{jd}}(K_3)$ of keyboard was 10, which indicated that among 28 groups of experts, 19 groups of the first batch had higher scores than those of the third batch; the value of $E_{\mathrm{tb}}(K_2)$ o tyre was 16, which indicated that among 32 groups of experts, 24 groups of the first batch had higher scores than those of the second batch; the value of $E_{\mathrm{tb}}(K_3)$ of tyre was 22, which indicated that among 32 groups of experts, 27 groups of the first batch had higher scores than those of the third batch; the values of $E(K_i)$ of keyboard and tyre are bigger than 0.

Based on the definition of $\mathrm{E}(K_i)$ and the above analysis, we find that the purchase decision quality of the first batch testees is bigger than the purchase decision quality of the second and third batches of testees. Thereafter, the experiment verifies H2: To some extent the information cocoon makes the consumer decision quality get worse.

5. Conclusions and outlook

5.1. Theoretical contributions

This paper has the following theoretical contributions: first, by reviewing the antecedents of information cocoons, this paper extends the studies on information cocoons. The current studies on information cocoons are centered on their antecedents and negative effects (Chernev et al., 2015; Johnson et al., 2013). Few scholars discuss the impact of information

Table 9
Statistical table for product $E(K_i)$.

Keyboard	Value	Tyre	Value
$\mathrm{E_{jd}}(K_2)$	18	$\mathrm{E_{tb}}(K_2)$	16
$\mathrm{E_{jd}}(K_3)$	10	$\mathrm{E_{tb}}(K_3)$	22

cocoons on consumer decision quality, which are strengthened by the AI recommendation from online purchase. The online purchase may produce a large amount of information. Based on this information, AI recommendation analyses consumers' preferences and continuously push the information that is in line with consumers' preferences to the consumers (Fleder et al., 2011). As such, the information cocoons are formed and strengthened. This paper applies the studies of information cocoons to the consumer online purchase, which provides a solid theoretical and empirical basis for studying the impact of information cocoons on consumer decision quality.

5.2. Practical significance

Based on the information browsed by consumers, AI recommendation can analyze consumers' preferences and therefore recommend the product information to consumers. Moreover, with consumers browsing more and more online products, the ability of AI to recommend is becoming stronger. It makes consumers trapped in the information cocoons and can't help it. Most consumers are passively using these online platforms. As such, we suggest that consumers change their preferences by asking others' opinions on their own preferences and actively searching for purchase information. The change in consumer preference will increase the chaos of shopping information. It also increases the types of recommended items, which will weaken the intensity of information cocoons. Hence, in the process of shopping, consumers can consciously pay attention to the similarity of the items recommended, roughly judge the strength of the information cocoon, consciously increase the degree of information chaos, and therefore increase the choices of products.

5.3. Limitations and outlook

This study has some limitations. First of all, the testees of the study are college students from Jiangxi University of Finance and economics. The experimental platforms are Jingdong.com and taobao.com. Though we randomly select the testees and both Jingdong are representative online shopping platforms in China, the conclusions of this study still need more experimental tests with various people and different shopping platforms. The follow-up studies can be designed by selecting the testees with different occupations and ages. Also, they can choose Suning, Amazon and other shopping platforms to conduct the experiment.

Second, this study has two contrast groups and they are G1 and G2; G1 and G3. We find that AI recommendation with account logging and without account logging respectively has a different impact on information cocoons. This is an interesting finding and worthy of the follow-up study. Comparatively, for AI recommendation with account logging, the constriction of information cocoons is relatively weak. This is perhaps because AI recommendation more depends on the current browsing information of consumers rather than consumers' historical browsing information and consumers' basic information such as gender, age and so forth.

5.4. Conclusions

For consumers, AI recommendation technology has penetrated into all aspects of online shopping. However, few studies are devoted to the relationship between AI recommendation, information cocoons and consumer decision quality. In this study, based on reviewing relevant literature, we adopt the method of group experiment and Delphi method to investigate the relationship of AI recommendation, consumers' preferences, information cocoons, and consumer decision quality. Based on the group experiment, we find: AI recommendation moderates the relationship of consumers' preferences and information cocoons. That is, given the consumers' preferences, the higher the degree of AI recommendation, the stronger the constriction of information cocoon. Moreover, this paper innovatively applies Delphi method to the evaluation of consumer decision quality. We find that the stronger the constriction of information cocoon, the lower the consumer decision quality. This study provides a new line of thinking for consumers to avoid falling into the information cocoons when shopping online. Besides, it quantitatively studies the constriction strength of information cocoons and the consumer decision quality for the first time, which opens up new ideas for the follow-up research. As such, this study has practical and theoretical significance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is supported by the NSFC (72002088); Jiangxi Natural Science Fund (20192BAB207010); Science and Technology Project Founded by the Education Department of Jiangxi Province (GJJ190292); Graduate Innovative Special Fund Projects of Jiangxi Province (YC2019-B090; YC2020-B138); Teaching Reform Research Project of Jiangxi Province Degree and Postgraduate Education (JXYJG-2020-085).

Appendices

Table A.1
Information Entropy and Gini impurity of testees on Jingdong Platform

ID	Information entropy	Gini impurity	ID	Information entropy	Gini impurity	ID	Information entropy	Gini impurity
101	6.08	0.98	201	2.12	0.84	301	2.79	0.20
102	4.44	0.92	202	3.14	0.85	302	3.71	0.90
103	2.70	0.77	203	3.85	0.93	303	3.49	0.89
104	4.52	0.96	204	2.78	0.80	304	3.23	0.89
105	5.26	0.95	205	2.32	0.66	305	3.28	0.85
106	4.69	0.91	206	2.22	0.74	306	3.11	0.73
107	3.95	0.88	207	2.35	0.73	307	4.49	0.88
108	3.71	0.95	208	2.13	0.79	308	3.10	0.53
109	5.47	0.92	209	2.69	0.92	309	4.01	0.68
110	4.17	0.96	210	3.64	0.59	310	2.31	0.89
111	5.60	0.82	211	1.85	0.77	311	3.59	0.90
112	2.92	0.95	212	2.84	0.73	312	3.69	0.75
113	5.23	0.93	213	2.22	0.88	313	2.46	0.90
114	5.04	0.85	214	3.00	0.80	314	3.43	0.68
115	3.23	0.85	215	2.74	0.83	315	4.58	0.72
116	3.46	0.94	216	2.58	0.86	316	4.71	0.96
117	4.45	0.84	217	3.09	0.86	317	5.12	0.79
118	3.69	0.89	218	2.81	0.88	318	4.04	0.81
119	3.84	0.93	219	3.10	0.89	319	2.97	0.88
120	4.26	0.94	220	3.17	0.82	320	3.10	0.95
121	5.06	0.88	221	2.52	0.71	321	4.65	0.80
122	4.04	0.90	222	2.33	0.71	322	3.79	0.78
123	4.12	0.89	223	2.40	0.79	323	2.66	0.88
124	4.28	0.93	224	2.74	0.84	324	3.12	0.67
125	4.61	0.83	225	3.18	0.88	325	1.92	0.95
126	4.12	0.95	226	2.00	0.84	326	2.86	0.81
127	3.41	0.94	227	3.43	0.75	327	4.93	0.85
128	4.80	0.94	228	2.75	0.83	328	2.82	0.86
T(G1)[a]	125.67	25.40	T(G2)[a]	78.76	22.51	T(G3)[a]	101.18	22.39

[a] In this experiment, G1 denotes the first batch of testees、G2 denotes the second batch of testees、G3 denotes the third batch of testees; T(G1)、T(G2)、T(G3) respectively denotes the total entropy and Gini impurity of G1、G2、G3 on Jingdong Platform.

Table A.2
Information Entropy and Gini impurity of testees on Taobao Platform

ID	Information entropy	Gini impurity	ID	Information entropy	Gini impurity	ID	Information entropy	Gini impurity
101	4.35	0.94	201	2.74	0.69	301	5.06	0.96
102	1.08	0.35	202	3.24	0.88	302	4.20	0.92
103	2.87	0.84	203	3.53	0.86	303	2.14	0.63
104	3.55	0.88	204	4.31	0.92	304	3.73	0.86
105	3.86	0.92	205	5.38	0.96	305	4.07	0.93
106	4.71	0.93	206	2.95	0.85	306	2.77	0.76
107	4.30	0.93	207	3.10	0.86	307	4.82	0.94
108	4.52	0.94	208	4.05	0.92	308	2.82	0.79
109	4.05	0.93	209	2.79	0.81	309	3.11	0.83
110	4.11	0.92	210	4.64	0.94	310	3.97	0.90
111	4.11	0.92	211	4.14	0.93	311	2.84	0.83
112	4.75	0.96	212	3.77	0.87	312	3.29	0.85
113	3.56	0.86	213	3.33	0.87	313	3.58	0.87
114	4.63	0.95	214	4.75	0.95	314	3.10	0.80
115	3.97	0.90	215	5.04	0.96	315	2.34	0.71
116	3.30	0.87	216	4.98	0.95	316	3.01	0.86
117	5.12	0.96	217	3.04	0.79	317	3.84	0.93
118	4.66	0.94	218	3.30	0.85	318	3.80	0.90
119	4.41	0.93	219	4.46	0.92	319	3.86	0.89
120	2.89	0.84	220	3.11	0.83	320	4.70	0.94
121	4.01	0.92	221	4.49	0.93	321	3.24	0.86

(*continued on next page*)

Table A.2 (*continued*)

ID	Information entropy	Gini impurity	ID	Information entropy	Gini impurity	ID	Information entropy	Gini impurity
122	3.43	0.88	222	2.93	0.77	322	5.84	0.97
123	3.01	0.86	223	3.68	0.86	323	4.54	0.94
124	4.10	0.92	224	3.51	0.89	324	4.00	0.92
125	3.66	0.89	225	3.12	0.84	325	2.50	0.79
126	2.70	0.80	226	3.26	0.84	326	2.97	0.83
127	3.69	0.89	227	2.54	0.70	327	4.12	0.90
128	3.79	0.90	228	2.66	0.74	328	4.49	0.92
129	3.94	0.91	229	3.32	0.86	329	3.04	0.82
130	2.20	0.69	230	4.31	0.92	330	3.48	0.88
131	3.65	0.90	231	3.66	0.87	331	3.39	0.85
132	4.21	0.93	232	2.82	0.79	332	4.73	0.94
T(G1)[a]	121.19	28.19	T(G2)[a]	116.96	27.63	T(G3)[a]	117.38	27.73

[a] In this experiment, G1 denotes the first batch of testees、G2 denotes the second batch of testees、G3 denotes the third batch of testees; T(G1)、T(G2)、T(G3) respectively denotes the total entropy and Gini impurity of G1、G2、G3 on Taobao Platform.

References

Aksoy, L., Cooil, B., & Lurie, N. H. (2011). Decision quality measures in recommendation agents research. *J. Interact. Market., 25*(2), 110–122.

Benbasat, L., & Nault, B. R. (1990). An evaluation of empirical research in managerial support systems. *Decis. Support Syst., 6*(3), 203–226.

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics Inf. Technol., 15*, 209–227.

Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: a conceptual review and meta-analysis. *J. Consum. Psychol., 25*(2), 333–358.

Dabholkar, P. A. (2006). Factors influencing consumer choice of a 'rating web site': an experimental investigation of an online interactive decision aid. *J. Market. Theor. Pract., 14*(4), 259–273.

Dellaert, B. G. C., Baker, T., & Johnson, E. J. (2017). Partitioning Sorted Sets: Overcoming Choice Overload while Maintaining Decision Quality. *Columbia Business School Research Paper.* https://doi.org/10.2139/ssrn.3082254

Fleder, D., Hosanagar, K., & Buja, A. (2011). Recommender systems and their effects on consumers: the fragmentation debate. *MCIS 2011 Proceedings, 59*.

Foster, R. (2012). *News Plurality in a Digital World*. Oxford: Reuters Institute for the Study of Journalism.

Guess, A., Lyons, B., Nyhan, B., & Reifler, J. (2018). Avoiding the Echo Chamber about Echo Chambers: Why Selective Exposure to Like-Minded Political News Is Less Prevalent than You Think. https://www.researchgate.net/publication/330144926_Avoiding_the_echo_chamber_about_echo_chambers_Why_selective_exposure_to_like-minded_political_news_is_less_prevalent_than_you_think.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: can one desire too much of a good thing? *J. Pers. Soc. Psychol., 79*, 995–1006.

Johnson, E. J., Hassin, R., Baker, T., Bajge, A. T., & Treuer, G. (2013). Can consumers make affordable care affordable? The value of choice architecture. *PloS One*. https://doi.org/10.1371/journal.pone.0081521

Mahajan, V., Green, P. E., & Goldberg, S. M. (1982). A conjoint model for measuring self- and cross-price/demand relationships. *J. Mar. Res., 19*, 334–342.

Negroponte, N. (1995). *Being Digital*. New York: Alfred A. Knopf.

Page, A. L., & Rosenbaum, H. F. (1987). Redesigning product lines with conjoint analysis: how sunbeam does it. *J. Prod. Innovat. Manag., 4*(2), 120–137.

Rust, R. T. (2001). The rise of E-service. *J. Serv. Res., 3*(4), 283–284.

Safran, M., & Che, D. (2017). Real-time recommendation algorithms for crowdsourcing systems. *Appl. Computing and Inform., 13*, 47–56.

Schulz, W., Dreyer, S., & Hagemeier, S. (2012). *Machtverschiebung in der öffentlichen Kommunikation*. Bonn: Friedrich-Ebert Stiftung.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell Syst. Tech. J., 27*(4), 623–656.

Sunstein, C. R. (2006). *Infotopia: How Many Minds Produce Knowledge*. Oxford: Oxford University Press.

Tian, B., Gu, T., & Lu, Y. (2019). Research on consumers' response to personalized recommendation avoidance in B2C E-business under filter bubble phenomenon. In *Proceedings of the 2019 International Conference on E-Business and E-Commerce Engineering*. https://doi.org/10.1145/3385061.3385068

Todd, P., & Benbasat, I. (1992). The use of information in decision making: an experimental investigation of the impact of computer-based decision aids. *MIS Q., 16*(3), 373–393.

Webster, J. (2010). User information regimes: how social media shape patterns of consumption. *Northwest. Univ. Law Rev., 104*, 593–612.

West, P. M., Ariely, D., Bellman, S., Bradlow, E., Huber, J., Johnson, E., Kahn, B., Little, J., & Schkade, D. (1999). Agents to the rescue? *Market. Lett., 10*(3), 285–300.

Xiao, B., & Benbasat, I. (2018). An empirical examination of the influence of biased personalized product recommendations on consumers' decision making outcomes. *Decis. Support Syst., 110*, 46–57.

Xiao, B., & Benbasat, I. (2007). E-commerce product recommendation agents: use, characteristics, and impact. *MIS Q., 31*(1), 137–209.

Xu, H., Chen, Z., Li, R., & Wang, C. (2020). *The Geometry of Information Cocoon: Analyzing the Cultural Space with Word Embedding Models*. New York: Cornell University.

Zeithaml, V. N., Bolton, N. R., Deighton, J., Keiningham, T. L., Lemon, K. N., & Petersen, J. A. (2006). Forward-looking focus: can firms have adaptive foresight? *J. Serv. Res., 9*(2), 168–183.

Zhao, Z. (2017). The characteristics and path analysis on public opinion spread of public emergency events in social media age. *International Conference on Contemporary Education, Social Sciences and Humanities, 124*, 1005–1007.

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? *Internet Policy Rev, 5*(1), 1–16.