Personality-Based Personalization of Online Store Features Using Genetic Programming: Analysis and Experiment

Alireza Kazeminia[1], Marjan Kaedi[2] and Beenazir Ganji[3]

[1] University of Isfahan, Faculty of Computer Engineering, Isfahan, Iran, kazeminia@eng.ui.ac.ir
[2] University of Isfahan, Faculty of Computer Engineering, Isfahan, Iran, kaedi@eng.ui.ac.ir
[3] Payame Noor University (PNU), Department of Computer Science, Tehran, Iran, b_ganji@pnu.ac.ir

Abstract

The decisions made by the customers in online environments are influenced by their personality characteristics. Each customer in an online environment relies more heavily on certain features of a store to make decisions while ignoring others. Thus, personalizing these features may streamline the decision-making process and increase satisfaction. In this paper, an intelligent method for personalizing the features of an online store according to the users' personality is presented. In the proposed method, using genetic programming several equations are developed to estimate how users with different personality characteristics prefer various features of an online store. These equations are then used for personalization of the store features to increase customers' satisfaction and persuade them to make larger purchases. The evaluation on a sample of 194 individuals indicates that the obtained equations are able to estimate the user's preferences with over 80% accuracy in most cases. In addition, empirical assessment of the obtained equations shows that the proposed personalization method improves the user satisfaction.

Keywords: Personalization, Online shopping, Personality, Decision-making style, Genetic Programming

1   Introduction

Each individual approaches decision-making differently: some make decisions instantly whereas others need time to consider various alternatives; some people rely on their intuitions to make decisions while others prefer to discuss their choices with friends. As with traditional shopping, customers in online stores behave differently while making decisions [15], [21]. In other words, they rely on different features of the store to make a decision [20]. Personalizing the features to each user's decision-making style can reduce the cognitive load of each decision, making the interaction more satisfying [31]. Personalization, as a strategy, has been shown to have several positive outcomes including customer satisfaction and increased sales [31]. In this paper, we propose an approach to estimating user preferences according to the decision-making styles; our results can be used for personalization in online stores.

Arguably, each customer in an online shopping context relies more heavily on certain features of a store to make decisions while paying less attention to others. In the retail industry, the one-size-fits-all approach, has become less and less effective over the year [26], [32]-[33]. Computers, and especially artificial intelligence, allow us to consider each person individually and personalize the interactions according to their behaviors, desires, attitudes, and beliefs [36]. Thus, it is possible to adapt the sales method to the customer's decision-making style in order to improve the online shopping experience.

This paper aims to model how customers with different personality traits and decision-making styles interact with an online store. It is assumed that the store provides adequate information so that the user can make an informed decision. Accordingly, several equations are created to estimate how much each customer relies on a certain feature or attribute to make a decision. The equations take the user decision-making style and extraversion as input and yield a value showing how important a feature is to the customer. Provided that the obtained values are sufficiently accurate, they may be used to personalize customer-store interactions.

The remainder of this paper is organized as follows. In Section 2, the underlying concepts and relevant previous work are reviewed. Section 3 presents our methodology. The proposed approach is detailed in Section 4 followed by the empirical assessment in Section 5. Next, a discussion of the results is presented in Section 6. Finally, concluding remarks are given in Section 7.

2   Background

This section is organized as follows: First, the effect of individual differences on the decision-making styles are described. Then, the existing tools and features in online stores are discussed. Finally, previous studies that explore the relation between individual differences and user behavior in online shopping environments are reviewed.

2.1   Individual Differences and Decision-Making

Personality is the cause of numerous differences among individuals. Research has shown that it may be used to explain differences in office and bedroom organization [25], favorite music [35], behavior in online social networks [55], information seeking behavior [28], and decision-making [1]. Every day, one is faced with hundreds of decisions, making decision-making an extremely important activity in people's lives.

Up to now, there have been various attempts to measure how individuals approach decisions. Examples include the General Decision-Making Style (GDMS) [18], Thinking and Working Style (TWS) [48], the Rational-Experiential Inventory (REI) [43], Intuitive-Analytical Judgment Style [53], and Cognitive Style Index (CSI) [2].

The GDMS distinguishes among five decision styles: (1) an *avoidant* style tries to postpone decisions; (2) a *dependent* style seeks advice from others; (3) an *intuitive* style believes in feelings and instincts; (3) a *rational* style requires a concise and logical evaluation of alternatives; finally, (5) a *spontaneous* style prefers to complete the decision-making process as fast as possible. The questionnaire is composed of five items for each style. Answers are given on a Likert scale of 1 (never) to 5 (always). Therefore, each participant receives a score between 5 and 25 on each scale. The maximal score determines the person's dominant style of decision-making. The instrument has been studied in different populations and exhibits adequate psychometric properties [22], [54]. Thunholm [56] uses the GDMS to conclude that the decision-making styles represent both habits and thinking practices. More recently, the instrument has been adapted to the context of healthcare to examine patient choice situations [18].

The TWS consists of ten items measuring *systematic* and *intuitive* styles. The items are similar to those of the GDMS. However, the former is more general and context-free as opposed to the latter which emphasizes decision-making [49]. Research suggests that the instruments correlate positively with respect to the systematic style [49]. Similarly, the REI is used to measure a person's *preference for the rational and experiential (faith in intuition)* decision-making [9]. Interestingly, males have been found to have higher degrees of Need for Cognition (NfC) [27], [41] while women usually score higher on faith in intuition [49].

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

As the name suggests, the Intuitive-Analytical Judgment Style [53] identifies two ways in which individuals respond to decision-making situations. The intuitive style prefers to base decisions on feelings whereas the analytical style tends to rely on calculations and comparisons. Finally, the CSI [2] also takes a similar approach. The instrument has been studied in organizational contexts. For instance, intuitive leaders have been shown to be more nurturing and less dominating while gaining more respect compared to their analytical counterparts [3].

From among the mentioned instruments, the GDMS is applied in the current study to determine the decision styles of online store users. The details are presented in Section 3. The reason of this choice is that GDMS has a relatively small number of items. In addition, it shows adequate psychometric properties and a reliable translation in the language of respondents is available.

In addition to models and questionnaires that consider the decision-making and judgment, there are numerous other models focusing on various dimensions of personality, some of which are related to decision-making. The Big Five model [14] is the most accepted and popular model of human personality [36] which identifies five domains or dimensions: Openness to experience, Consciousness, Extraversion, Agreeableness, and Neuroticism. Each domain is composed of 12 facets. Research on human personality has shown that decision-making is influenced by personality [11]; for instance, extraverts tend to exhibit higher degrees of toughness in their decisions [17]. Based on these studies, the extraversion dimension of the Big Five is employed in the current study for describing the behavior of online store users (Section 3).

2.2   User Interaction in Online Store

As mentioned earlier, each customer in an online environment tends to rely more heavily on certain features of a store to make decisions while paying less attention to others. In fact, various aspects of the shopping experience may be personalized to the customer's decision-making style in an attempt to increase the likelihood that the customer will make a purchase [34]. Thus, one may adapt the features of an online store such that the user is able to make decisions according to his/her own decision-making style.

Online stores typically offer the users a host of information and features. Each feature serves a specific purpose. For instance, the comparison tool allows the users to compare the details of several products. Research suggests that these tools may be used by the customers to decrease their cognitive effort [39]. Despite the variety, the features can be divided into three broad categories, as follows [37]:

- Content: These include text, images, product descriptions, labels, ratings, user comments, and expert reviews.

- User Interface: This category can be further divided into appearance (fonts, colors, and patterns) and tools (comparison capability, number of products in each page, sort criteria, user comments, ratings, and capability to share on online social networks). Note that the tools allowing the users to rate and comment on products belong to this category; the actual comments and ratings, however, comprise content.

- Additional Offers: These are used to entice customers to make larger purchases. They can be subdivided into related products (such as accessories) and complementary offers (e.g. insurance for the purchased product).

Given the variety of tools and features in different online stores, in this study, we select seven features that are more commonly found in an online store and cover all three of the abovementioned categories. This is because online shopping is considered in a general sense and so the selected features are common and exist in nearly all online stores. Furthermore, the choices are guided by extant literature indicating the relevance of different features. Each of the features may have several states or values. For example, comments by other users may be shown or hidden or the comparison feature may be displayed in a way that attracts the user attention and prompts usage. Table 1 presents the selected features along with possible values for each one.

2.3   Previous Works

The link between user behavior in online shopping environments and individual differences has been studied from a variety of perspectives. Nevertheless, existing literature may be divided into two general categories: (1) studies that explore the online decision-making process in a general sense, or with the objective to persuade customers; and (2) studies that examine how personality relates to customer intentions and behaviors. In the following, we briefly review related studies in both categories.

The customer decision-making process is often viewed as a linear process consisting of five different stages: need recognition, information search, evaluation of alternatives, purchase, and post-purchase [51]. Recently, Karimi et al. [34] proposed a framework containing loops to represent the non-linear nature of decision-making in the real world.

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

Yet, individuals tend to display different behaviors in performing the process; these differences form different decision-making styles.

Table 1: Selected features and possible values

Category	Feature	Value(s)	References
Content	Expert Reviews	Shown vs. Hidden	[29], [58]
Content	Product Description	Appeal towards Extraversion vs. Introversion	[12]
User Interface	Compare	Prominent vs. Normal	[39], [57]
User Interface	User Comments	Shown vs. Hidden	[4], [58]
User Interface	Product Ratings	General vs. Detailed	[4], [16], [19]
Additional Offers	Recommended Products	Products, such as accessories, that complement the chosen product.	[37]
Additional Offers	Complementary Offers	Insurance and Extended Warranty	[37]

Karimi et al. [34] found that online shopping decision behaviors can be adequately explained using two constructs of decision-making style and prior product knowledge. Maximizers and satisfiers with the same levels of knowledge behave very differently. Taking a different approach, Rezaei [46] identified six different shopping styles and their respective preferred attributes in an online store. It was shown that various sets of retailer attributes are considered important for different styles. For instance, recreational shoppers *shop for the pleasure of shopping*. However, other price-conscious shoppers are interested in the lowest price while another group of shoppers value uniqueness and visual appeal in an online store. Another instrument known as the Consumer Decision-Making Style (CDMS) has been used to divide online and offline customers into different segments [46].

According to the study of Oinas-Kukkonen and Harjumaa [42], development of a persuasive system includes three phases: first, understanding the essential issues behind the persuasive system, second, analyzing the context for persuasive systems to identify the intents, events, and strategies for using such system, and finally, designing the new features for a new system or evaluating the features of an existing system. The Elaboration Likelihood Model (ELM) of persuasion [45] has been applied to the online shopping context. More specifically, it has been used to study customer purchase intentions. According to Yang, in the presence of a positive peripheral cue, elaboration leads to higher purchase intentions; however, a negative peripheral cue seems to counteract the effects of elaboration [13]. Moreover, the central route is more appropriate for agreeable and conscientious users with utilitarian purchase intentions whereas the peripheral route is recommended for extraverted and open individuals who have hedonic intentions [12]. Ashraf and Narongsak [6] found similar results using the concept of regulatory fit which is closely related to the ELM.

Kaptein [31] demonstrated that by taking individual differences into account conversion rates can be increased by up to 40 percent. Using the social influence strategies, the author showed that each person is more susceptible to a single strategy [32]. In another study, these differences were found to be stable both over time and context [31]. They have been used to generate adaptive persuasive messages for selling books and music [31], reducing snack consumption [33], and taking the stairs [33].

The bulk of personality research in online shopping environments relies on the Big-Five model. Tsao and Chang [18] concluded that extraversion, neuroticism, and openness to experience influence hedonic purchase motivations. Customers with higher degrees of neuroticism, agreeableness, or openness to experience behave in a utility-motivated manner, which leads them to use the search and comparison capabilities more frequently. Furthermore, agreeableness, extroversion, openness to experience, and emotional stability have been positively linked with online fashion shopping [47]. Bosnjak et al. [10] presented a hierarchy of personality traits that online purchase decisions. Their results indicate that in addition to the Big Five traits, NfC, need for material resources, and need for arousal predict users' intentions to shop online. The Myers-Briggs Type Indicator (MBTI) has also been used to confirm that individual differences influence how individuals make shopping decisions in online environments [7]. Sanayei et al. [50] used the MBTI to study online shopping intentions. Their results indicate that introverted, sensing, and thinking individuals are more likely to engage in online shopping.

3 Methodology

In order to study customer decision-making behavior in online shopping, one needs to know how different customers rely on various features in an online store [44]. Preferences may be extracted either explicitly or implicitly. Explicit

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

methods directly inquire about user preferences whereas implicit methods depend on user behavior [37]. For the sake of feasibility, in this paper we use an explicit instrument i.e. a survey to collect user preferences.

A questionnaire composed of 46 items in two sections was used for data collection purposes. Demographic questions on gender and age were included at the end of the questionnaire. The first section included the GDMS (25 items) [18] together with the Extraversion items from the Big Five model (12 items) [14].

Two factors contribute to the choice of GDMS: (1) the instrument has a relatively small number of ítems, which is likely to enhance users' willingness to participate in the study; and (2) since the study was conducted in Iran, irrespective of the choice, the instrument had to be in Persian; however, the authors were not able to find reliable Persian translations of any other instruments nor does a native inventory of decision-making style exist. The translated version of the GDMS was obtained from the office of Vice Dean of Research at the Faculty of Education and Psychology in University of Isfahan. Furthermore, based on previous studies pointing to the effects of extraversion on decision-making, we employed the extraversion scale of the Big Five, obtained from the same source, alongside the GDMS in an attempt to better understand online shopping behavior.

Although it seems intuitive to claim that the styles identified by the GDMS manifest in online shopping environments, to the best of our knowledge, the GDMS has not been used in this context previously.

The second section of the questionnaire aimed to find out about user preferences. Thus, the authors developed a total of nine items that inquired about user preferences regarding the features identified in Table 1: two for Content (Product Description and Content), four pertaining to User Interface (Compare, Comments, and Ratings), and three for Additional Offers. In other words, participants were asked about their preferences in using each of the seven features in Table 1. Answers were given on a Likert scale of 1 (never) to 5 (always). The items (translated into English) can be seen in Table 2.

A note on the validity and reliability of the second section is presented here. While validity refers to the extent to *evidence and theory support the interpretation of test scores for proposed uses of the test*, reliability is defined as the tendency towards consistency found in repeated measurements of the same phenomenon [5]. Based on these definitions, it seems unnecessary to study the reliability of the items as they directly inquire about user preferences, which are relatively enduring over longer periods of time. Moreover, given the ample evidence in previous works (Section 2), a statistical examination of item validity was not essential. Nevertheless, face validity was checked by two external reviewers and the ítems were adjusted according to their feedback.

Using these answers, it is possible to have a general picture of how the participant interacts with an online store. Note that, the features are regarded as independent of how they are implemented. As mentioned earlier, the instrument covers several generic features that exist in almost any online store. Furthermore, the participants indicate how frequently they rely on a specific feature in order to make a more definite decision. Thus, it is possible to argue that, when making a decision, the user is very likely to seek his/her intended feature regardless of how it is implemented. For instance, a user with confidence in expert reviews is probably not concerned about the presentation of the review, as long as sufficient information is given.

The population of the study includes the individuals with previous online shopping experience. A convenience sample of 194 individuals studying at the University of Isfahan was selected to participate in the study. The university is among the most prestigious educational institutes in the country. The questionnaire was self-administered in groups at the end of the lectures under the supervision of the authors. The participants were asked to fill out a questionnaire consisting of the aforementioned 46 items. The responses to the questionnaire took nearly 15 minutes.

Table 2: Items from the second section of the study questionnaire

No.	Item	Never	Rarely	Sometimes	Often	Always
1	Product descriptions give me the necessary information to make a decision.					
2	I read the comments other users leave for different products.					
3	I consider general product ratings important.					
4	Instead of a single score, I prefer detailed ratings.					
5	In order to choose among different products, I compare the details.					
6	I normally read product reviews by experts.					
7	I buy accessories that go along with the product I purchase.					
8	I prefer to buy the accessories as bundles.					
9	I'm willing to make extra payments to reduce the risks associated with my purchase.					

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

The participants were relatively young individuals with Persian cultural backgrounds. It was made sure that each individual had previously made several online purchases (digital devices, clothing, household items, etc.) and was sufficiently familiar with popular Iranian online stores. With respect to age, the participants were between 19 and 38 years old and nearly 56 percent were male. Twenty-eight participants gave incomplete or imprecise answers which were removed from the dataset. Therefore, 166 records were used in the analysis phase.

The results are analyzed in two steps. Firstly, correlation analysis is performed to determine how the variables relate to each other. Secondly, using the significant relationships identified in the previous step, a number of mathematical equations are developed in order to estimate user preferences based on their personality. Later on, the equations may be applied to the new users in order to estimate their preferences by knowing their decision-making styles. Thus, the need to explicitly ask for user preferences in eliminated. In addition, instead of using questionnaire to extract the users decision-making styles, the decision-making style of a new user can be estimated using the implicit methods [23]-[24], [30] by tracking his/her behavior.

The estimation task is essentially an optimization problem wherein an equation with a sufficiently small error term links one or more independent variables to a dependent variable. Evolutionary algorithms are well-suited to optimization problems and provide good approximate solutions. In particular, Genetic Algorithms (GAs) have been quite successful in efficiently solving optimization problems [40].

A GA mimics the process of evolution in biology. Hypotheses are described as bit strings or chromosomes, whose merit is evaluated according to a fitness function. GAs start with an initial collection of chromosomes known, as the population, and successively generates offspring with higher fitness values. Two operations are involved in the process of creating generations: crossover and mutation [40].

The equations, in this paper, are obtained using Genetic Programming (GP), which is a variant of GA. In GP, instead of bit strings, the evolving population is composed of computer programs. Thus, it may be used to generate computer programs and algorithms which are shown by trees. Each function call is represented as a tree node, with the inputs as its branches. Although optimality is not guaranteed, genetic programs offer exceptional performance in many cases [40].

4   Analysis and Results

This section focuses on analyzing the data and reports the results. In order to improve accuracy, four outlier responses were identified and subsequently eliminated using the nearest neighborhood method [5]. Correlation analysis was conducted on the remaining 65 records to identify significant relationships between personality variables and user preferences for each feature (Subsection 4.1). Thereafter, several mathematical equations were developed based on the identified significant relationships. The purpose of these equations is to estimate user preferences according to the extraversion and decision-making style (Subsection 4.2).

4.1   Correlation Analysis

The first step in analyzing the data involves calculating correlation coefficients so that significant relationships can be identified. Two sets of variables, i.e. independent and dependent, are considered as follows:

- Dependent variables include the five decision-making styles: Avoidant (A), Dependent (D), Intuitive (I), Rational (R), and Spontaneous (S), as well as the personality dimension of Extraversion (Ext).

- For each item in the second section of the questionnaire, one dependent variable is considered which determines user preference.

The analysis aims to determine significant relationships among the variables so that mathematical equations can be formed to estimate user preferences. Significant relationships were identified according to their correlation coefficients at levels $p < 0.1$ and $p < 0.05$ for 65 records [58], yielding 0.247 and 0.288 respectively. The calculated coefficients for all pairs of variables are shown in Table 3.

4.2   Mathematical Equations

In the previous step, the correlations among the variables were calculated to reveal significant relationships. In this step, we develop several equations that aim to estimate user preferences for each of the nine features on a discrete scale of 1 to 5. The significance of the equations lies in their ability to approximate how the new users, whose decision-making styles are known, are likely to behave in an online store. This opens up a host of possibilities for personalizing user interactions (which is outside the scope of this paper).

Personality-Based Personalization of Online Store Features Using Genetic
Programming: Analysis and Experiment

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

Table 3: Correlation coefficients

No.	Variable	A	D	I	R	S	Ext	Number of significant relationships at level of p < 0.05	at level of p < 0.1
1	Product description	0.252*	0.132	0.053	-0.05	0.263*	0.051	0	2 (A, S)
2	User comments	0.085	-0.010	0.251*	0.223	0.229	0.111	0	1 (I)
3	General product ratings	0.079	0.021	0.150	0.074	0.304**	0.131	1 (S)	1 (S)
4	Detailed product ratings	-0.210	-0.08	0.084	0.301**	0.061	0.222	1 (R)	1 (R)
5	Compare	-0.190	0.164	0.231	0.487**	-0.120	0.223	1 (R)	1 (R)
6	Expert reviews	0.006	0.311**	0.170	0.409**	0.094	0.245	2 (D, R)	2 (D, R)
7	Accessories	-0.260*	-0.09	0.201	0.200	0.047	0.255*	0	2 (A, Ext)
8	Accessory bundle	0.061	0.166	-0.240	0.221	-0.030	0.110	0	0
9	Risk reduction	-0.160	-0.060	0.278*	0.305**	0.146	0.207	1 (R)	2 (I, R)
	Average	-0.047	0.03	0.092	0.133	0.061	0.162	0	0

* p < 0.1
** p < 0.05
Number of significant relationships at the level of p < 0.05: 6
Number of significant relationships at the level of p < 0.1: 12

To construct the equations, the first set of variables, i.e. the extraversion and decision-making style, act as independent variables, while user preferences are considered dependent. In order to develop each of the nine equations, only the variables with significant relationships (i.e. p<0.05 and p<0.1) are considered.

As mentioned before, the equations are constructed using GP. In doing so, we used matrix laboratory (MATLAB) software. The parameters are as follows. Maximum tree depth was set to nine to prevent bloat (i.e. overgrown trees with little contribution to fitness). The limit was applied to tree depth as opposed to the number of non-terminal nodes. The trees were initialized using the Ramped Half-and-Half method which results in a diverse population of both balanced and unbalanced trees having various depths [52]. Creating the population requires functions and terminals. In our experiments, the addition, subtraction, multiplication, and natural logarithm operators were selected. The constant values obtained from the participants, in addition to random numbers between 0 and 1, were used as the terminal nodes.

Fitness was calculated as the sum of the absolute difference between the expected output value and the returned value. Smaller differences were preferred. Each case was run for 500 generations. In GP, the crossover operator selects random branches from two parent trees and swaps the respective branches to create two offspring. The mutation operator, on the other hand, works by choosing a random node and substituting it with a different function or terminal node [40]. Once new individuals were created using either of the operators, those exceeding the maximum allowed depth (i.e. nine) were rejected. In each generation, parent trees were selected according to lexicographic parsimony pressure wherein the parent individual is designated as the best of a randomly selected group of individuals and shorter trees are preferred [38]. The size of the randomly selected group of individuals was set to one percent of the entire population. In GP, the likelihood of using each operator depends on its performance: an operator that exhibits better performance is assigned higher probability values. However, operators without children since the last adaptation, experience a considerable increase in their probability so that they have the chance to reproduce. The values 0.25 and 0.01 were used as the higher and lower bounds for the probabilities.

The collected dataset is partitioned into training data and test data. Seventy percent of the records (i.e., 116 records) are randomly selected and considered as the training data and the remaining records (i.e., 50 records) are considered as test data. To conduct the cross validation, the data partitioning process repeats for 10 times and the averaged Root Mean Square Error (RMSE) and the best obtained equations over the 10 times of cross validation are reported in Table 4. The GP tree constructed for item 2 is depicted in Figure 1 as an example. In this figure, *minus*, *plus*, and *mult* represent the subtraction, addition, and multiplication operators, respectively and I indicates the intuitive decision-making style.

To investigate the effectiveness of the proposed method, the results of the proposed method are compared to those of regression analysis. For this purpose, GP is replaced with the multiple linear regression analysis in the described

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

cross validation process. The averaged RMSE of multiple linear regression analysis and the best obtained equations over the 10 times of cross validation are reported in Table 5. By comparing Tables 4 and 5, it is concluded that the RMSE obtained from the regression analysis is much higher than that of our proposed method. Thus, the complicated equations obtained from the proposed method improve the average accuracy of predictions.

Table 4: The best mathematical equations and the averaged error values obtained from the proposed method

| No.	Variable	Significant Relationship		Equation	RMSE (Training set)	RMSE (Test set)
		p < 0.05	p < 0.1			
1	Product description		*A, S*	$\ln(A^2 + s\ln(S) + 0.002A + 4.2\,S + 0.27)$	0.167	0.192
2	User comments		*I*	$0.0021\,I^3 + 0.82\,I^2 - 2.87$	0.172	0.185
3	General product ratings	*S*		$\ln(S) + 0.21\,S + 1.93$	0.192	0.203
4	Detailed product ratings	*R*		$0.031\,R^2 + 0.04\,\ln(R) + 8.61$	0.158	0.181
5	Compare	*R*		$2.3\,\ln(R) + 0.003\,R - 1.18$	0.179	0.195
6	Expert reviews	*D, R*		$0.78\,\ln(2.53\,D^2) + 0.9\,\ln(R^2) + 0.07$	0.206	0.214
7	Accessories		*A, Ext*	$\ln(0.003\,A + 0.68\ln(Ext))$	0.198	0.211
8	Accessory bundle	-	-	-	-	-
9	Risk reduction	*R*	*I*	$0.87\,\ln(I^2 + 1.13\,R)$	0.205	0.218
Average:					0.184	0.199
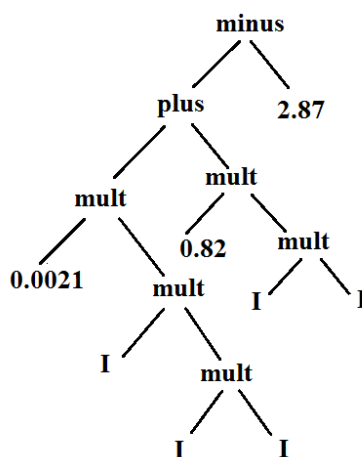
Figure 1: The GP tree of item 2

As mentioned, the test data applied for evaluation was obtained from the Iranian students of University of Isfahan. These students were between 19 and 38 years old. In addition to the mentioned evaluation, it is required to evaluate the obtained equations on the dataset with other features in order to determine whether the precision of our proposed method is influenced by factors like culture, education, and age of the users. To study and assess this issue, the data related to personality features, decision-making style, and preferences of the users are collected from the following three groups:

- Group 1: 35 foreign students of ages within 19-38 from China, Iraq, Syria and Afghanistan enrolled in University of Isfahan

- Group 2: 35 Iranian people above age 40 with higher education

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

- Group 3: 35 Iranian people of ages within 19-38 with no higher education

The questionnaire is distributed among the three groups. The equations presented in Table 4 are applied on the collected data and the error rate of equations in predicting the participants' preferences is assessed. The outcome is presented in Table 6.

Table 5: The best mathematical equations and the averaged error values obtained from the multiple linear regression analysis

No.	Variable	Significant Relationship		Equation	RMSE (Training set)	RMSE (Test set)
		p < 0.05	p < 0.1			
1	Product description		*A, S*	$3.6039 + 0.1013\,A + 0.0452\,S$	0.254	0.323
2	User comments		*I*	$-2.5350 + 0.4411\,I$	0.247	0.296
3	General product ratings	*S*		$1.6768 + 0.1535\,S$	0.353	0.401
4	Detailed product ratings	*R*		$-4.9714 + 0.5254\,R$	0.277	0.376
5	Compare	*R*		$3.0310 + 0.1371\,R$	0.228	0.267
6	Expert reviews	*D, R*		$7.4056 + 0.1006\,D + 0.1056\,R$	0.312	0.388
7	Accessories		*A, Ext*	$0.2457 + 0.0009\,A + 0.0055\,Ext$	0.331	0.396
8	Accessory bundle	-	-	-	-	-
9	Risk reduction	*R*	*I*	$3.2852 + 0.0030\,R + 0.0959\,I$	0.286	0.351
Average:					0.286	0.349

Table 6: The averaged error values obtained from the proposed method on the test data of groups 1, 2, and 3

No.	Variable	Evaluation on group 1		Evaluation on group 2		Evaluation on group 3	
		RMSE	t-test result	RMSE	t-test result	RMSE	t-test result
1	Product description	0.216	N	0.199	N	0.197	N
2	User comments	0.231	N	0.192	N	0.189	N
3	General product ratings	0.286	S	0.209	N	0.199	N
4	Detailed product ratings	0.281	S	0.192	N	0.184	N
5	Compare	0.313	S	0.206	N	0.197	N
6	Expert reviews	0.317	S	0.219	N	0.223	S
7	Accessories	0.289	N	0.225	S	0.218	N
8	Accessory bundle	-	-	-	-	-	-
9	Risk reduction	0.341	S	0.227	N	0.223	N
Average:		0.284		0.208		0.203	

By comparing Tables 5 and 6 it is deduced that the RSME measure for the data of the groups 2 and 3 are very close to that of test data reported in Table 5. The one-tailed t-test at the significance level of 0.05 reveals that there exists no significant difference among these results in the majority of cases. On the other hand, the RMSE of group 1 is

higher than that of test data presented in Table 5. The one-tailed t-test verifies the significance of difference in the majority of cases. It could be deduced that the equations that map the user personality features and decision-making style to his/her preferences are not influenced by education level and age of users, but by their nationality and the cultural backgrounds.

Consequently, to personalize an online store according to the proposed method it is necessary to collect the training data from the target population of that online-store, in order to extract the equations that fit for the given culture.

5   Empirical Assessment

In this section, the proposed method is implemented in practice and is evaluated by a number of users. For this purpose, a prototype online store has been developed with all the features needed for this research. Thirty three products including digital devices and clothing have been registered in this online store. In order to prevent the impact of external factors, the products used are all within the same price range and dummy brand names are chosen for them. In addition, an identical image was used for all products. Forty four users volunteered to evaluate this system. These people were between the ages of 19 and 38, and 50% of them were males. These users, without being aware, were randomly divided into experimental and control groups with each group consisting of 22 users. Members of the experimental group worked with an online store that according to the proposed method of the study, featured personalization. But for the control group users, no personalization was done. Each user first completes a questionnaire containing 37 items when registering on the system.

As mentioned in Section 3, this questionnaire includes the GDMS (25 items) [55] together with the Extraversion items from the Big Five model (12 items) [49].

The values of the equations in Table 4 are in the range of 1 to 5, while the values of the features of Table 1 are binary. For example, the values of *shown* and *hidden* values for the *Expert Reviews* feature are represented by 0 and 1, respectively. To match the differences between these values, the value of 3 is considered as the threshold. Thus, the values 4 and 5 obtained from the equations are considered as the first value of each of the features of Table 1 and the values of 1, 2, and 3 are mapped to the second value of the features in this table. Users were asked to complete the registration process, search the products and select their desired products. Finally, after the users worked with the system, they were asked to complete a questionnaire that includes three questions to measure *customer satisfaction from interacting with the store*, *satisfaction with decisions taken* and *willingness to return to the store*. Participants responded numerically in the range of 1 (completely opposite) to 5 (fully agree). Table 7 shows the average results of each question for the experimental and control groups. The *t*-test results indicate whether the difference between the experimental group and control group is Significant (S) or Non-significant (N). The following results are obtained from the table:

- The degree of satisfaction of the experimental group from interacting with the system is greater than the control group's satisfaction from interacting with the system.

- The degree of satisfaction of the experimental group with decisions taken is greater than the control group's satisfaction with decisions taken.

- The willingness of users to return to the store in the experimental group has no significant difference with the control group.

Table 7: The results of the customers satisfaction and their willingness to return to the store for the experimental group and control group

	Averaged on experimental group members	Averaged on control group members	Degree of improvement	Percentage of improvement	Result of one-tailed *t*-test (S/N)
Satisfaction from interacting with the store	4.27	3.83	0.44	11.48 %	S
Satisfaction with decisions taken	4.31	4.05	0.26	6.41 %	S
Willingness to return to the store	3.87	3.84	0.03	0.78 %	N

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

Furthermore, we applied the one-tailed t-test at the significance level of 0.05 to investigate whether or not there is any statistically significant difference between the results of the experimental group and control group. The results demonstrate that the satisfaction of the experimental group is significantly greater than the control group's satisfaction. However, there is no significant difference between the *willingness to return to the store* of the members of experimental group and control group.

6   Discussion

As seen in Tables 3 and 4, the rational style is the most frequently significant variable and appears in four of the equations. This may be due to the fact that shopping in online environments is more about utility than pleasure [42], making the approach more attractive to more rational individuals. Furthermore, only three of the five styles (i.e. dependent, rational, and spontaneous) are strongly correlated with a portion of the examined user preferences. General product ratings were found to influence spontaneous individuals. This is expected as the style is characterized by a desire to make a decision as fast as possible [18]; thus, the ratings may serve as a suitable heuristic. In contrast, rational individuals tend to place emphasis on ratings which distinguish among various aspects of the product. Intuitively, the dependent user seeks advice from experts by reading available reviews. Finally, it is interesting to note that, extraversion is only weakly correlated with the preference for accessories. This is somehow consistent with previous findings in that the trait is associated with hedonic shopping behaviors [42]. However, the weak correlation may require further validation in future studies.

According to the results, user preferences can be estimated with an accuracy of approximately 80 percent. More importantly, the equations are simple and only need one or two input variables. Thus, each one can be calculated with speed. The (quadratic or cubic) polynomial or logarithmic equations may be readily implemented using almost any technology. In order to estimate user preferences for one feature, it is not necessary to consider the entire set of independent variables. For instance, using only the rational style, user preferences for several features may be determined. Furthermore, the error values are acceptable; since preferences are graded on a discrete scale of 1 to 5, an error of 20% indicates a one-step error. For example, very high preference (i.e. 5) may be confused with high preference (i.e. 4). Using adaptive machine learning algorithms, it is possible to reduce the estimation error. However, this is outside the scope of this paper.

Based on our results, it is clear that certain features in online stores are more favored by certain groups of users. The implications of these results are four-fold: Firstly, the equations suggest a pathway toward personalizing the user interfaces. In other words, knowing the users' decision-making styles, it is possible to personalize the user interface of an online store so that it better matches the user expectations. For instance, dependent people were found to have confidence in what experts think of a product; thus, it is better to emphasize the feature in the user interface to ease the decision-making process.

Secondly, as seen in Table 4, rationality appears to be the dominant approach to decision-making in the context of online shopping. This fact must be taken into account to guide the design of more effective user interactions.

Thirdly, by comparing Tables 4 and 6 it is observed that the extracted equations are influences by the participants' cultural background. Therefore, it is necessary to apply training data regarding target populations' cultures to extract the equations.

Finally, although the notion of personalizing the user interface is not a novel idea, the problem has rarely been approached from a mathematical perspective. The equations developed in this paper allow user preferences to be estimated instantly and eliminate the need for capturing and storing complex data about each user. Additionally, the estimated values may serve as a starting point of the onboarding process for the new users so that when the new users visit the store for the first time adequate personalization can be applied, provided that the user decision-making style is known.

7   Conclusion

In this paper, a method for estimating user preferences based on personality and decision-making style was proposed. To this end, a survey with 46 items was conducted on a random sample of 194 individuals. Following the elimination of incomplete or imprecise responses as well as outliers, correlations among the variables were calculated and the significant relationships were identified. These relationships were used to develop mathematical equations using MATLAB. The analysis and empirical evaluation of the obtained equations show that, knowing the users' decision-making styles, it is possible to personalize the user interface of an online store so that it increases the users' satisfaction. According to our study, rationality is the dominant decision-making style in the context of online shopping and it must be taken into account for effective personalization of the user interface of online stores. This paper used a questionnaire to ask user preferences explicitly. However, users are generally unwilling to answer such questions with sufficient care; thus, adequate accuracy may be hard to obtain.

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

Overall, our results serve as a first step toward personalizing the online shopping experience through a mathematical approach. However, there are a large number of issues that need to be addressed before these results can be applied in the real-world circumstances. One important aspect is to extend the proposed method for the international online stores which have customers with different cultural backgrounds. For this purpose, the training data should be collected from people with different cultural norms. The frequency distribution of cultural backgrounds in the training data should be similar to that of online store target population. Besides, the explicit methods can be replaced by implicit methods to estimate the users' personality [23]-[24], [30].

Future works need to focus on analyzing implicit extraction of user personality and decision-making style in online stores. In addition, the users' privacy concerns [8] about the personalized stores should be studied. Finally, other dimensions of personality as well as other online store features may be explored.

References

[1]   Y. Akbulut-Bailey, C. A. Looney and R. S. Poston, Personalization systems for decisions involving risk: the interactive effects of user personality and suggestive guidance, Information Technology Management, vol. XXIV, no. 4, pp. 34-50, 2013.

[2]   W. Allinson and J. Hayes, The cognitive style index: a measure of intuition-analysis for organizational research, Journal of Management Studies, vol. 33, no. 1, pp. 119-135,1996.

[3]   W. Allinson, S. J. Armstrong and J. Hayes, The effects of cognitive style on leader-member exchange: a study of manager-subordinate dyads, Journal of Occupational and Organizational Psychology, vol. 74, no. 2, pp. 201-220, 2001.

[4]   E. Almonani, W. Husain, O. Y. San, A. Almomani, and M. Al-Betar, Mobile game approach to prevent childhood obesity using persuasive technology, in Proceedings 2014 International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, Malaysia, 2014, pp. 1-5.

[5]   American Educational Research Association, American Psychological Association and National Council on Measurement in Education, Standards for educational and psychological testing. Washington, DC: American Educational Research Association, 2014.

[6]   R. Ashraf and N. Thongpapanl, Connecting with and converting shoppers into customers: investigating the role of regulatory fit in the online customer's decision-making process, Journal of Interactive Marketing, vol. 32, pp. 13-25, 2015.

[7]   R. Barkhi and L. Wallace, The impact of personality type on purchasing decisions in virtual stores, Information Technology and Management, vol. 8, no. 4, pp. 313-330, 2007.

[8]   B. Berendt, O. Günther and S. Spiekermann, Privacy in e-commerce: stated preferences vs. actual behavior, Communications of the ACM, vol. 48, no. 4, pp. 101-106, 2005.

[9]   F. Björklund and M. Bäckström, Individual differences in processing styles: validity of the rational–experiential inventory, Scandinavian Journal of Psychology, vol. 49, no. 5, pp. 439-446, 2008.

[10]   M. Bosnjak, M. Galesic and T. Tuten, Personality determinants of online shopping: explaining online purchase intentions using a hierarchical approach, Journal of Business Research, vol. 60, no. 6, pp. 597-605, 2007.

[11]   K. A. Byrne, C. D. Silasi-Mansat and D. A. Worthy, Who chokes under pressure? The Big Five personality traits and decision-making under pressure, Personality and Individual Differences, vol. 74, pp. 22-28, 2015.

[12]   S. H. Chen and K. P. Lee, The role of personality traits and perceived values in persuasion: an elaboration likelihood model perspective on online shopping, Social Behavior and Personality: an International Journal, vol. 36, pp. 1379-1399, 2008.

[13]   Y.-X. Chen, S.-S. Chiang, S.-Y. Chih, W.-C. Liao, S.-Y. Lin, S.-H. Yang, et al., Opportunities for persuasive technology to motivate heavy computer users for stretching exercise, in Proceedings of the 9th International Conference on Persuasive Technology, Padua, Italy, 2014, pp. 25-30.

[14]   P. T. Costa Jr, R. R. McCrae and D. A. Dye, Facet scales for agreeableness and conscientiousness: a revision of the neo personality inventory, Personality and Individual Differences, vol. 12, no. 9, pp. 887-898, 1991.

[15]   S. Dakduk, E. ter Horst, Z. Santalla, G. Molina, and J. Malavé, Customer behavior in electronic commerce: a bayesian approach, Journal of Theoretical and Applied Electronic Commerce Research, vol. 12, no. 2, pp. 1-20, 2017.

[16]   B. De Langhe, P. M. Fernbach and D. R. Lichtenstein, Navigating by the stars: Investigating the actual and perceived validity of online user ratings, Journal of Consumer Research, vol. 42, no. 6, pp. 817-833, 2016.

[17]   D. Farkas and G. Orosz, The link between ego-resiliency and changes in Big Five traits after decision making: The case of extraversion, Personality and Individual Differences, vol. 55, no. 4, pp. 440-445, 2013.

[18]   S. Fischer, K. Soyez and S. Gurtner, Adapting scott and bruce's general decision-making style inventory to patient decision making in provider choice, Med Decis Making, vol. 35, pp. 525-32, 2015.

[19]   J. Flanagin, M. J. Metzger, R. Pure, A. Markov, and E. Hartsell, Mitigating risk in ecommerce transactions: perceptions of information credibility and the role of user-generated ratings in product quality and purchase intention, Electronic Commerce Research, vol. 14, pp. 1-23, 2014.

[20]   J. Fogel and S. Zachariah, Intentions to use the yelp review website and purchase behavior after reading reviews, Journal of Theoretical and Applied Electronic Commerce Research, vol. 12, pp. 53-67, 2017.

[21]   B. Fogg, A behavior model for persuasive design, in Proceedings of the 4th International Conference on Persuasive Technology, Claremont, California, 2009, pp. 1-7.

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

[22] E. Gambetti, M. Fabbri, L. Bensi, and L. Tonetti, A contribution to the Italian validation of the General Decision-making Style Inventory, Personality and Individual Differences, vol. 44, pp. 842-852, 2008.

[23] I. Golkar Amnieh and M. Kaedi, Using estimated personality of social network members for finding influential nodes in viral marketing, Cybernetics and Systems, vol. 46, no. 5, pp. 355-378, 2015.

[24] I. Golkar and M. Kaedi, Developing a model for estimating the extraversion degree of social network members using the information extracted from the graph structure, Journal of Modeling in Engineering, vol. 13, no. 43, pp. 91-106, 2016.

[25] S. D. Gosling, S. J. Ko, T. Mannarelli, M. E. Morris, A room with a cue: Personality judgments based on offices and bedrooms, Personality and Social Psychology, vol. 82, pp. 379-398, 2002.

[26] A. Goy, L. Ardissono and G. Petrone, Personalization in e-commerce applications, in The Adaptive Web (P. Brusilovsky, A. Kobsa, W. Nejdl, Eds.). Berlin: Springer, 2007, pp. 485-520.

[27] P. Haugtvedt, R. E. Petty, and J. T. Cacioppo, Need for cognition and advertising: understanding the role of personality variables in consumer behavior, Journal of Consumer Psychology, vol. 1, pp. 239-260, 1992.

[28] J. Heinström. (2003, October). Five personality dimensions and their influence on information behaviour, Information Research. [online]. vol. 9, no. 1, paper 165, Available : http://InformationR.net/ir/9-1/paper165.html.

[29] N. I. Holleschovsky and E. Constantinides, Impact of online product reviews on purchasing decisions, in Proceedings WEBIST 2016: The 12th International Conference on Web Information Systems and Technologies, Rome, 2016, pp. 271-278.

[30] M. Kaedi, N. Ghasem-Aghaee, T. I. Ören and L. PourMohammadBagher, Anger evaluation for fuzzy agents with dynamic personality, Mathematical and Computer Modelling of Dynamical Systems, vol. 15, no. 6, pp. 535-553, 2009.

[31] M. Kaptein, P. Markopoulos, B. de Ruyter, and E. Aarts, Personalised persuasion in ambient intelligence, Journal of Ambient Intelligence and Humanized Computing, vol.1, no.1,pp. 43-56, 2010.

[32] M. Kaptein, P. Markopoulos, B. de Ruyter, and E. Aarts, Can you be persuaded? individual differences in susceptibility to persuasion, in Human-Computer Interaction–INTERACT 2009, vol. 5729, (T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R.O. Prates, M. Winckler). Berlin: Springer, 2009, pp. 115-118.

[33] M. Kaptein, Building persuasion profiles in the wild: using mobile devices as identifiers, presented at the ACM Conference on Human Factors in Computing Systems (CHI), Vancouver, May 7 - 12, 2011.

[34] S. Karimi, K. N. Papamichail and C. P. Holland, The effect of prior knowledge and decision-making style on the online purchase decision-making process: A typology of consumer shopping behaviour, Decision Support Systems, vol. 77, pp. 137-147, 2015.

[35] M. Kopacz, Personality and music preferences: The influence of personality traits on preferences regarding musical elements, Journal of Music Therapy, vol. 42, pp. 216-239, 2005.

[36] M. Kosinski, Y. Bachrach, P. Kohli, D. Stillwell, and T. Graepel, Manifestations of user personality in website choice and behaviour on online social networks, Machine Learning, vol. 95, pp. 357-380, 2014.

[37] K. Kwon and C. Kim, How to design personalization in a context of customer retention: Who personalizes what and to what extent?, Electronic Commerce Research and Applications, vol. 11, pp. 101-116, 2012.

[38] S. Luke and L. Panait, Lexicographic parsimony pressure, in Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '02), Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2001, pp. 829-836.

[39] N. H. Lurie and N. Wen, Simple decision aids and consumer decision making, Journal of Retailing, vol. 90, pp. 511-523, 2014.

[40] T. M. Mitchell, Machine Learning. New York: McGraw-Hill, Inc., 1997.

[41] S. C. Nettelhorst and L. A. Brannon, The effect of advertisement choice, sex, and need for cognition on attention, Computers in Human Behavior, vol. 28, pp. 1315-1320, 2012.

[42] H. Oinas-Kukkonen and M. Harjumaa, Persuasive systems design: Key issues, process model, and system features, Communications of the Association for Information Systems, vol. 24, no. 1, 2009.

[43] R. Pacini and S. Epstein, The relation of rational and experiential information processing styles to personality, basic beliefs, and the ratio-bias phenomenon, Journal of Personality and Social Psychology, vol. 76, pp. 972-87, 1999.

[44] P. Papatla, Do online shopping styles affect preferred site attributes? An empirical investigation and retailing implications, Journal of Retailing and Consumer Services, vol. 18, pp. 362-369, 2011.

[45] R. E. Petty and J. T. Cacioppo, The Elaboration Likelihood Model of Persuasion, in Advances in Experimental Social Psychology. vol. 19, (B. Leonard, Ed.). Cambridge, Massachusetts: Academic Press, 1986, pp. 123-205.

[46] S. Rezaei, Segmenting consumer decision-making styles (CDMS) toward marketing practice: A partial least squares (PLS) path modeling approach, Journal of Retailing and Consumer Services, vol. 22, pp. 1-15, 2015.

[47] S. Roy, R. Sethuraman and R. Saran, The effect of demographic and personality characteristics on fashion shopping proneness, International Journal of Retail & Distribution Management, vol. 44, pp. 426-447, 2016.

[48] L. Sagiv, S. Arieli, J. Goldenberg and A. Goldschmidt, Structure and freedom in creativity: The interplay between externally imposed structure and personal cognitive style, Journal of Organizational Behavior, vol. 31, pp. 1086-1110, 2010.

[49] L. Sagiv, A. Amit, D. Ein-Gar, and S. Arieli, Not All Great Minds Think Alike: Systematic and Intuitive Cognitive Styles, Journal of Personality, vol. 82, pp. 402-417, 2014.

[50] A. Sanayei, N. A. Bazargan and A. Ansari, The impact of introversion/extroversion on online shopping intention (Case study: Computer and cell phone accessories), in Proceedgins 2016 10th International Conference on e-Commerce in Developing Countries: with focus on e-Tourism (ECDC), Tehran, Iran, 2016, pp. 1-7.

Alireza Kazeminia
Marjan Kaedi
Beenazir Ganii

[51] S. Senecal, P. J. Kalczynski and J. Nantel, Consumers' decision-making process and their online shopping behavior: a clickstream analysis, Journal of Business Research, vol. 58, pp. 1599-1608, 2005.
[52] S. Silva. (2016, March) GPLAB - a genetic programming toolbox for MATLAB. GPLAB. [Online]. Available: http://gplab.sourceforge.net/index.html.
[53] L. Sjöberg, Intuitive vs. analytical decision making: which is preferred?, Scandinavian Journal of Management, vol. 19, pp. 17-29, 2003.
[54] D. P. Spicer and E. Sadler-Smith, An examination of the general decision making style questionnaire in two UK samples, Journal of Managerial Psychology, vol. 20, pp. 137-149, 2005.
[55] J. Staiano, B. Lepri, N. Aharony, F. Pianesi, N. Sebe, and A. Pentland, Friends don't lie: inferring personality traits from social network structure, in Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, Pennsylvania, 2012, pp. 321-330.
[56] P. Thunholm, Decision-making style: habit, style or both?, Personality and Individual Differences, vol. 36, pp. 931-944, 2004.
[57] W.-C. Tsao and H.-R. Chang, Exploring the impact of personality traits on online shopping behavior, African Journal of Business Managemen, vol. 4, pp. 313-330, 2010.
[58] Z. K. Zhang, S. J. Zhao, C. M. K. Cheung, and M. K. O. Lee, Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model, Decision Support Systems, vol. 67, pp. 78-89, 2014.