Information Systems Frontiers
https://doi.org/10.1007/s10796-020-09980-9

Personalize, Summarize or Let them Read? A Study on Online Word of Mouth Strategies and Consumer Decision Process

Mahesh Balan U[1] · Saji K. Mathew[1]

Abstract
This study integrates theories of task complexity and cognitive stopping rule to understand how complexity of information environment impacts uncertainty, effectiveness and efficiency of consumers' decision process. Using the reviews provided by an online retailer, we develop an e-commerce environment with three levels of complexity: high with raw textual reviews, medium with attribute-level review summaries and low with web personalization strategy based on attribute preferences extracted from online reviews. In a controlled lab experiment, users took buying decisions under different levels of complexity. Our analyses of clickstream data showed that users' effectiveness and efficiency were the highest in review based personalized environment. However, between groups who received summarized and textual reviews, the latter demonstrated apparently higher effectiveness and efficiency in decision making, which went against our anticipation. Further investigation showed that users simplified decision process when exposed to raw reviews. These results further inform reviews-based personalization strategy in e-commerce.

Keywords Online word of mouth · Summarized reviews · Web personalization · E-commerce · Complexity · Consumer behavior

1 Introduction

The online world has not only changed the way consumers shop for goods but also in the complexity involved in decision making. A consumer decides to buy or not to buy a product based on the cognitive processing of available information. This process is directly dependent on the extent of relevant information that a consumer has access to, to arrive at a final choice. In e-commerce setting the information can be in the form of numerical product ratings or textual product reviews. The potential of online product reviews in influencing consumer buying decisions and sales has been reported in prior studies (Archak et al. 2011; Ghose and Ipeirotis 2007). However, the multiplicity of alternatives and degree of uncertainty with online Word of Mouth (WoM) have generated a new challenge to the consumer in allocating attentional resources (Kahneman 1973a; Moray 1967). For example, a quick review of the top 50 mobile phones and tablets (based on sales rank) in Amazon.in yielded on an average 3050 reviews online. Furthermore, this information is very dynamic in nature and is subject to updating with user generated content by future buyers. In this context, how could online product reviews be made available to online consumers in content and format that would attract the attention of an online consumer and provide support in buying decisions? Further, how would such manipulation of online reviews impact a consumers' decision process?

The Internet is arguably the biggest revolution of the modern world. Its meteoric growth has provided a plethora of opportunities for new businesses and new business models for existing businesses. The information system that connects the buyers with the sellers and facilitates the transactions between them is the core engine that drives the electronic marketplace. In contrast to a brick and mortar store, the electronic marketplace has more access to information about products in the form of online word-of-mouth and has lesser search costs by showing different types of products in a single place. The online word-of-mouth comprises of information such as consumer product reviews, consumer product ratings, response to consumer FAQs and helpfulness scores of the reviews; of these, consumer reviews are most rich in attribute-level

✉ Mahesh Balan U
u.maheshbalan@gmail.com

Saji K. Mathew
saji@iitm.ac.in

[1] Department of Management Studies, Indian Institute of Technology Madras, Chennai, India

Published online: 28 January 2020

information (Archak et al. 2011; Balan and Mathew 2016b; Liu and Karahanna, 2017).

Consumer product reviews are analogous to reading Shakespeare, due to its complex nature. The complexity is primarily due to (i) Information Overload (ii) Multi-dimensionality (iii) Attribute-richness and (iv) Degree of information conflict (Liu and Karahanna, 2017). The information overload refers to the quantum of reviews, whereas multi-dimensionality refers to the usage of multiple languages in a single review text, and also different synonyms for the same word. The reviews generally discuss about multiple attributes of a product and it adds more attribute richness to it. Furthermore, consumer product ratings which are numerical in nature follow a bimodal distribution (Ghose and Ipeirotis 2010; Hu et al. 2009; Hu et al. 2006), and the same holds true for the text sentiments in product reviews as well. In most of the products, almost equal number of people opine their reviews in conflicting sentiments. This higher degree of information conflict makes the task of reading reviews a challenging task.

Consumer product reviews are textual in nature and they give a 360-degree overview of a product experience (Archak et al. 2011). Notwithstanding the helpfulness of online WoM in consumer's decision-making process, the availability of a plethora of textual information increases the number of alternatives and their levels for a decision maker, which adds to task complexity (Campbell 1988). The volume of information available in an e-commerce platform includes product information from the manufacturer, consumer reviews, consumer ratings and product information from the e-commerce platform.

Following Campbell (1988), preferential choice in online environment can be understood as a decision task where the decision maker has to exercise choice decisions on a multi-alternative, multi-level situation. Consumer reviews provide several attributes and users' qualitative ratings of the levels of these attributes. While performing complex tasks for preferential choice, individuals utilize different heuristics that will keep the information processing demands of the situation within the bounds of their limited mental capacity (Payne 1976). If the amount of information processing exceeds a certain limit, people's cognitive capacity will likely influence the learning performance (Norman and Bobrow 1975). In IS research multi-channel presentation has been proposed to address task complexity (Jiang and Benbasat 2007). Another stream of research in e-commerce has investigated the impact of web personalization on consumers' information processing and decision outcomes (Ho et al. 2011; Krishnaraju et al. 2016; Tam and Ho 2006). Since users' attentional resources are limited (Kahneman 1973b), presenting relevant content to the user is likely to reduce task complexity and cognitive efforts required in information processing. A widely used strategy to provide relevant content is Web Personalization, and prior research has empirically demonstrated that web-personalization influences the decision outcomes of the user (Das et al. 2007; Ho et al. 2011; Krishnaraju et al. 2016; Tam and Ho 2006).

In the backdrop of online buyers' complex information processing environment, we seek to understand how personalized review feedback (personalization based on review summaries) alters information environment and influences decision making process of an online consumer. In particular, we aim to study how a personalized environment based on consumer reviews relates to decision making uncertainty, effectiveness (in terms of acquiring accurate information) and efficiency (in terms of not wasting time on unnecessary information) of the consumers' online decision process. Our research integrates theory of task complexity and cognitive stopping rule and extends web personalization to online product reviews to understand how online consumers' decision process varies with e-commerce environments posing different levels of task complexity. We build upon the works of Archak et al. (2011) who evaluated the pricing power of consumer reviews, Tam and Ho (2006) who evaluated the effect of personalization on consumer's cognitive processes while making a buying decision and the work of Browne and Pitts (2004) who studied stopping rule usage in information search.

This paper is organized as follows: in Section 2, 3, 4 and 5 we give a review of literature and theory development, followed by hypotheses development and methodology in Sections 5 and 6. We present analyses subsequently in Section 7 and inferences are discussed in in Section 8 which also includes results from a focus group discussion. The theoretical and managerial implications are discussed in Sections 9 and 10, followed by future scope for research and conclusion in Sections 11 and 12.

2 Online Word of Mouth

Online word of mouth has received a distinct attention in information systems research (Archak et al. 2011; Clemons et al. 2006; Forman et al. 2008; Gao et al. 2015; Ghose and Ipeirotis 2010; Hu et al. 2017; Liu and Karahanna 2017). Our review showed four important focus areas pertaining to online word-of-mouth - impact of online word-of-mouth on sales, reviewer behavior and bias, dimensionality of reviews and economic impact of reviews which is also termed as economining (Balan and Mathew 2015). Other related areas of research include task complexity (Campbell 1988; Jiang and Benbasat 2007), personalization (Ho and Bodoff 2014; Ho et al. 2011; Tam and Ho 2006) and social cognition (Wyer Jr and Srull 2014). Our work integrates concepts from these areas to develop understanding about personalization of reviews and consumer decision process.

Extending the definition of WoM given by Arndt (1967), online word of mouth can be defined as a message from the person who bought a product *online* about her or his experience to a person who is interested in buying the product, where the person who is giving the review has a non-commercial intention. Several studies have analyzed the effect of consumer reviews on sales performance in various segments like books (Chevalier and Mayzlin 2006), movies (Liu 2006) and professional services (Gao et al. 2015).

Attribute summarization techniques of consumer reviews have also been addressed in prior literature which involve text summarization (Zhuang et al. 2006), sentiment summarization (Beineke et al. 2004) and aspect level summary from reviews (Blair-Goldensohn et al. 2008; Balan and Mathew 2016a, b). Techniques such as sentiment analysis, part-of-speech tagging and named entity recognition have been employed to extract meaning from online consumer reviews. Consumer reviews are unstructured in nature, huge in volume, serial in presentation and hence consumers underutilize them due to their limited search capabilities. This brings a need for visualization of summary for consumers like opinion blocks (Alper et al. 2011). Blair-Goldensohn et al. (2008) extracted sentiment summary for relevant aspects of a restaurant like food, ambience, service and value. They tagged the reviews by these aspects and summarized their sentiments. In order to eliminate exaggerated reviews subjectivity score was also computed for the reviews and were used in filtering reviews.

3 Web Personalization

Web personalization refers to the process of adapting web content to meet the specific needs of users to maximize sales (Korper and Ellis 2001). Every consumer is unique and is looking for specific information online to make their buying decision. (Slovic 1995) showed that customers with the same preference follow the same process of preference elicitation, but the preferences elicited through these processes need not be the same for every customer. Drawing on the foundations of social cognition research, prior studies have conceptualized web personalization as consisting of two important constructs: self-reference and content relevance (Tam and Ho 2006; Wyer Jr and Srull 2014). Self-reference is the ability to process information related to the self distinctly from processing other information. The famous example is that of a cocktail-party, where an individual has the ability to distinguish her or his name being called despite the noise inside the party hall. Content relevance is the measure of relevance in the information available online that can be useful for the consumer.

Consumers find it really difficult to make their choices when there are multitudes of choices available (Iyengar and Lepper 2000; Balan and Mathew 2017). So, they may look for more information to justify that their choice is rational (Yaniv 2004). Jiang and Benbasat (2007) evaluated the cognitive efforts and complexity involved in online decision process of a consumer and two constructs emerged as determinants of the satisfaction of an online consumer: actual product knowledge available in the website and the perceived website diagnosticity, which is the extent to which the consumer believes the website provides sufficient information about the product to make the buying decision.

Tam and Ho (2006) studied the impact of web personalization on the decision-making process of online consumers. They conducted a multi-factor experiment to test the effect of self-reference, content relevance and goal specificity on the cognitive process of consumers while making a buying decision. This study reported that self-reference, content relevance and goal specificity influence cognitive processes, attention and purchase decisions of consumers. A similar study by Sheng et al. (2008) examined the impact of personalization and content relevance on the privacy concerns of the consumer. Ho et al. (2011) evaluated the impact of adaptive web personalization at various time points by conducting a multi-factor experiment with different types of recommender systems and different presentation timings. This study reported that the quality of the recommender system improved over time with more information available about the consumer, but the probability of the consumer accepting the recommended product decreased over time. Their findings established this by adopting models from consumer search theory and estimating the optimum timing to introduce personalization for products such as music and books. Krishnaraju et al. (2016) studied user acceptance of personalization in the context of e-government services and their findings suggested that personalizing based on self-reference and content-relevance had a significant moderating effect on the user's intention to use government services online in India. There are several studies on the impact of web stimuli like user interface of a website and user engagement by a website on user attention measured through number of clicks and purchase decision (Jiang and Benbasat 2007; Tam and Ho 2006). However, the impact of these stimuli in a personalized environment could be very different from that of a non-personalized environment. The differential impact of personalized and non-personalized product reviews is unexplored to the best of our knowledge. We address this gap in our research.

In Information Systems research, complexity has been studied with a focus on information presentation (Jiang and Benbasat 2007). In the context of e-commerce, there are a few studies which focus on the swaying effect of attribute preferences (Liu and Karahanna, 2017) and self-selection bias (Hu et al. 2017). Liu and Karahanna (2017) provided early evidence that attribute rich review information influences the attribute preference of the consumers. In a controlled laboratory experiment, they found that users' attribute preferences are swayed by the degree of attribute-richness. On the personalization front, there are more studies in Information Systems literature which focus on the decision-making behavior aspect (Tam and Ho 2005, 2006; Ho et al. 2011; Ho and Bodoff 2014). The complexity lens of personalization is not explored sufficiently in research on web personalization and the decision-making behavior is not studied except for self-selection bias (Hu et al. 2017) and swaying effect (Liu and Karahanna, 2017). We address this research gap by proposing a novel personalization using review information and also studying the influence of both review information and personalization on the decision making process of the consumer through the theoretical lenses of complexity and cognitive stopping rule.

4 Theory Development

4.1 Theory of Complexity

Research in task complexity dates back to early 1970s, particularly related to information-processing and decision-making literature. Scholars have analyzed task complexity both in terms of task characteristics and also the cognitive processes involved during task performance by a task-doer. According to Campbell (1988) task complexity can be classified into three – first, as a subjective psychological experience, second, as an interaction between the task and the task-doer and third as an objective aspect related to task. Among these, complexity as an interaction between the task and the task-doer has received more research attention emphasizing complexity of a task in terms of the task-doer (Shaw 1971) and multiplicity of paths to complete the task (Frost and Mahoney 1976). Non-prescribed tasks, which are characterized by incompletely defined alternatives and multiple ways to complete are generally found more complex than prescribed tasks, which are characterized by clearly defined alternatives.

When task complexity is understood as an objective phenomenon, a task can be more complex if the objective of the task could be accomplished in many ways where the task-doer has to find the optimal way to do it. In this scenario, the task is more complex because of the difficulty in comprehending information associated with the alternatives. Any objective that has information diversity, information loading or higher rate of information change will have higher complexity associated with it. There are four characteristics that explain these diversities in information– (a) multiple paths, (b) multiple outcomes, (c) interdependence among paths, and (d) uncertainties in paths and outcomes (Campbell 1988). Tasks primarily concerned with multiple outcomes has been termed as decision tasks and multiple outcomes based on the alternative chosen characterize this category of tasks. Online buying decisions based on review information could arguably belong to this category.

Task complexity is also related to cognitive effort. Factors like the mode of representation of information could influence task complexity and in turn cognitive effort. For example, graphical presentation of stock prices and consumer reviews in e-commerce could reduce cognitive efforts of the task-doer, thus decreasing complexity (Hammond 1986). Task complexity has also been conceptualized as a moderator of diagnosticity (Jiang and Benbasat 2007). They evaluated the cognitive efforts and complexity involved in the task of a consumer by comparing four different presentation formats in e-commerce. Two constructs emerged as determinants of the satisfaction of an online consumer: actual product knowledge available in the website and perceived website diagnosticity. The findings from this study suggest that perceived website diagnosticity influences the consumer more than actual product knowledge (Jiang and Benbasat 2007). Consumers develop different cognitive strategies to make their decisions and two dominant strategies involve decomposition and holistic approaches (Simon 1981; Browne et al. 2007). In a decomposition strategy, consumers give importance to every attribute of a product and the resulting mental representation of the product wherein every attribute can be individually recognized by the consumer. Whereas, in a holistic strategy consumers' mental representation is "organic integrative; the person represents the task as a whole and acts based on his 'sense' or 'image' or 'gist' of the situation rather than on individual elements" (Browne et al. 2007, p.93).

4.2 Cognitive Stopping Rule

Information search is required for all problem solving and decision-making tasks (Simon 1981). Information search is defined as the process of seeking knowledge or data about a problem, situation or artifact and the process is characterized by divergent thinking which involves the person opening her or his mind to new possibilities and perspectives (Couger 1996). Prior studies have shown evidence that people use heuristics or rule-of-thumb to terminate their information search process when exposed to information overload (Browne et al. 2007; Browne and Pitts 2004; Busemeyer and Rapoport 1988; Payne 1976).

A cognitive stopping rule is a heuristic or rule-of-thumb that a person invokes to terminate his or her information search process (Browne and Pitts 2004). Here decision making is understood as a multi-stage process which involves information acquisition for design of alternatives, evaluation of alternatives and the final decision making. The stopping rules used in all these stages are distinct (Browne and Pitts 2004). In the first stage, information is gathered by the decision maker from various sources till sufficient information is acquired and this judgement of information sufficiency is achieved by invoking a stopping rule or heuristic. The sufficiency of information is characterized by the correctness and completeness of the information (Smith et al. 1991). Once the decision maker believes to have sufficient information, she or he stops gathering information and moves to the next stage in the decision process which involves evaluation of alternatives. The heuristic or rule-of-thumb used by the decision maker to determine the completeness of information is called the cognitive stopping rule (Browne and Pitts 2004).

Browne and Pitts (2004) suggested 5 cognitive stopping rules in information requirements determination setting and later empirically investigated them (Browne et al. 2007). The five cognitive rules which are driven by the mode of information processing of the decision maker are mental list rule, representational stability rule, difference threshold rule, magnitude threshold rule and single criterion rule. In mental list rule, the person has a list of items to be satisfied before he or she stops collecting information. In contrast, representational stability rule suggests that the person will stop collecting information when the mental representation of her or his understanding stops shifting and gets stabilized. Difference threshold rule has an a priori difference level set by the decision maker and she or he stops collecting information when the learning level is lesser than the difference threshold set before. In magnitude threshold rule, the person stops searching information if the person reaches the cumulative amount of information threshold set by her or him before. In single criterion rule, as the name suggests, the person focuses on one criterion and she or he stops searching information if all relevant information about that criterion is collected.

Prior research has also showed that consumers search for information until the marginal value of additional information equals the marginal cost of acquiring that additional information (Stigler 1961). Here the loss from terminating information search in addition to considering the expected value of the additional information and expected cost to acquire it are evaluated together. Extending the concept of cognitive stopping rule to online context, consumers invoke heuristic or rule-of-thumb to terminate information search about products that they want to buy when exposed to information overload. Information overload in the online context can be in the form of online word-of-mouth like consumer product rating, manufacturer product description and consumer product reviews. We build on these concepts to develop our research hypotheses.

In sum, prior research has evaluated the dimensionality of reviews (Dellarocas et al. 2007; Kim and Hovy 2004; Liu et al. 2007; Zhang and Varadarajan 2006), the effect of reviews on sales (Chen et al. 2004; Chevalier and Mayzlin 2006; Forman et al. 2008; Sun 2012; Gao et al. 2015). (eg.: Archak et al. 2011; Chevalier and Mayzlin 2006) and also how presentation formats reduce task complexity (eg.: Jiang and Benbasat 2007); however, research that focuses on the intersection of these two areas i.e. understanding the complexity of review information and how review information could be used to reduce complexity of consumers' online decision process is scarce. In order to address this gap, this study proposes to develop a novel personalization approach based on consumer reviews and test how personalization based on reviews influences the uncertainty, effectiveness (in terms of acquiring accurate information) and efficiency (in terms of not wasting time on unnecessary information) of consumer decisions.

5 Hypotheses Development

Consumers in general have some beliefs about the product in their working memory while visiting an e-commerce platform to make their purchase (Archak et al. 2011). These beliefs follow a normal distribution and after receiving the information available in the website, the beliefs get updated to new mean values (Archak et al. 2011; Ho and Bodoff 2014). Consumers use the online word-of-mouth to make their buying decisions and invest some cognitive efforts to process these beliefs in their working memory while making a buying decision. However, due to the limited information processing capacity of the human brain and the multi-dimensionality of online word-of-mouth, the task of buying online becomes a complex task to the consumer. From the standpoint of objective task complexity, there are three primary properties of a complex task: (a) number of sources of information, (b) number of alternatives with each source of information, and (c) the degree of uncertainty in the information (Campbell 1988; Schroder et al. 1967). The online e-commerce platform has multiple sources of information like consumer product ratings, consumer product reviews, manufacturer's product information and e-commerce platform's product information. In addition to these, the consumers also have the opportunity to look at various review websites as an alternative source that provides additional information about the product and consumer feedback. The information present in consumer product ratings and consumer product reviews are very dynamic which get updated with the addition of every new review, thus the uncertainty in the information is very high. These attributes of online buying environment reflect properties of complexity and hence online buying can be termed as a complex task. Given a complex task with overload of voluminous information consumers use heuristics or rule-of-thumb, defined as stopping rule, while making their decisions (Browne et al. 2007; Browne and Pitts 2004; Busemeyer and Rapoport 1988; Payne 1976). Building on the theory of complexity and cognitive stopping rule, we argue that consumers' search behavior varies with the task complexity of online buying and this task complexity of buying is dependent upon the information environment.

5.1 Task Complexity and Information Environment

Task complexity is a function of the information processing complexity of the decision and higher the information processing requirements of a task, higher the task complexity (Campbell 1988; Byström and Järvelin 1995; Balan and Mathew 2016a, b). When a task demands high information processing requirements, then the decision maker experiences information overloading (Moore et al. 1996; Balan and Mathew 2016a, b) and this deteriorates the learning effect of the decision maker (Baddeley 1992). Uncertainty experienced by a user in this context is defined as the information gap which is the difference between the information required by a user to make a decision and the information that is present with the user to make a decision. Thus, higher the information gap, more the uncertainty. We use the number of changes in consumers' shopping cart before making an online buying decision as an indicator of uncertainty faced by the consumer since uncertainty is a measure of complexity (Campbell 1988).

Furthermore, when a decision maker faces information overload, decision behavior could follow a satisficing principle (Simon, 1955) where a decision maker would search the space of alternatives till the point an alternative that satisfies the minimum aspiration level of the decision maker is found. Elimination by aspects (EBA), another classical decision model proposed by Tversky (1972) describes choice as a covert elimination process wherein a decision maker follows a process of elimination from a candidate list of alternatives (Payne 1976, pp. 367–368). In an e-commerce setting, if a user is very confident with her or his decision, the certainty in decision process will be reflected by making less changes in her or his product cart. In contrast, if a user is in a state of information overload, she or he will add a product, then change the product when better products are encountered, and this elimination will be repeated till the user reaches a satisficing condition.

We thus posit that consumers in online environment would make number of changes in their shopping cart (an indicator of uncertainty faced by the consumer) based on the nature of information available to them in the form of product reviews. Reviews that are voluminous in raw form (free flowing text) present the highest level of complexity, reviews when summarized as scores presents medium complexity and reviews when summarized and personalized according to user preferences further reduces complexity (Tam and Ho 2006; Balan and Mathew 2016a, b). Personalization of summarized review not only reduces information overload, but also makes the information relevant to the online user.

H1a: Users who receive review summaries (medium complexity) will make lesser number of changes in the product cart to buy a product (thus experiencing lesser uncertainty), as compared to those who receive voluminous reviews in raw form (high complexity).

H1b: Users when exposed to personalized review feedback (low complexity) will make lesser number of changes in the product cart to buy a product (thus experiencing lesser uncertainty), as compared to those who receive review summaries (medium complexity).

5.2 Efficiency and Effectiveness of Decision Process

We further argue that consumers use different cognitive strategies, cognitive stopping rules and hence exhibit different effectiveness and efficiency in making their buying decisions for different levels of buying task complexity. In an online buying environment, consumers could have conflicting objectives of effectiveness and efficiency (Browne and Pitts 2004). They are subject to the risk of over acquisition or under acquisition of information. Over acquisition leads to holding too much information than necessary and hence incur more costs in terms of energy, time and cognitive efforts to process that information. In contrast, under acquisition involves acquiring less information than necessary leading to the risk of a decision error (Browne and Pitts 2004). Based on this dilemma, we further argue that an environment that bridges the gap between voluminous information of the e-commerce environment and the limited information processing capacity of the consumer could serve to balance effectiveness and efficiency. In particular, we expect that summarized reviews and personalized review feedback would lead to greater levels of efficiencies in online decision making.

H2a: Users when exposed to attribute-level review summaries (medium complexity) will require lesser time to make buying decision (higher decision-making efficiency), as compared to those who are exposed to voluminous reviews in raw form (high complexity).

H2b: Users when exposed to personalized review feedback (low complexity) will require lesser time to make buying decision (higher decision-making efficiency), as compared to those who are exposed to attribute-level review summaries (medium complexity).

Extending cognitive stopping rule to online environment, consumers sample products till they are confident that the sampled products are enough to choose the best product from. Stopping rule, which in mathematical terms (Appendix 1) is the point at which the marginal benefit to the consumer from sampling additional items is less than the marginal cost, in the online context is the rule that determines the point beyond which the person would stop searching and choose the best product from the sampled items (Ho and Bodoff 2014). Thus, stopping rule, in the online context, is a measure of the optimal number of products to be sampled by the consumer before she or he makes the buying decision. We expect that effectiveness of decision process will be reflected in the information seeking behavior of the consumer. We advance these concepts to argue our research hypotheses further.

H3a: Users when exposed to attribute-level review summaries (medium complexity) will seek lesser information about a product to make buying decision (higher decision-making effectiveness), as compared to those who are exposed to voluminous reviews in raw form (high complexity).

H3b: Users when exposed to personalized review feedback (low complexity) will seek lesser information about a product to make buying decision (higher decision-making effectiveness), as compared to those who are exposed to attribute-level review summaries (medium complexity).

This research proposes a new e-commerce environment which we believe will reduce the buying task complexity experienced by the user in terms of effectiveness (i.e acquiring relevant information) and efficiency (i.e. not wasting time on unnecessary information) by providing a visual summary of reviews and also attribute specific sentiment scores. We now describe the procedure we followed to extract sentiment scores of every attribute from the product reviews and provide them as a visual representation to the user. We believe that the new information provided in the form of sentiment scores for every attribute will enable the consumer to employ a decomposition strategy to make a mental representation of the product and hence reduce the buying task complexity encountered by the consumer.

6 Research Methodology

Our study adopted a sequential mixed-methods approach employing a multi-factor experimental method to test our hypotheses followed by a focused group discussion to further seek insights on our findings. According to (Eisenhardt 1989) mixed-methods approach provides a synergistic view of the evidence, fostering divergent perspectives to strengthen data collection and analyses. Since our findings would benefit from the narratives of the participants of the experiment in providing a rich set of explanations, our laboratory experiment was followed up with three focus group discussions. This approach has also been followed by IS scholars in previous research (Fitch 2007; Belanger 2012).

We conducted a multi-group laboratory-controlled experiment with three different groups exposed to three different levels of complexity. The first group of participants with high complexity had multiple sources of information, multiple alternatives available and very high uncertainty in the information, thus having high complexity (Campbell 1988). This group, which is the control group, has voluminous reviews in raw form, no personalization and the products were displayed randomly. The second group with medium complexity was exposed to all the attributes of the control group, but in addition to those, the participants also received an attribute-level review summary which showed the average sentiment score of every attribute from the reviews. This summarized information in the form of attribute-level sentiment scores reduced the need for the user to look for other sources of information and hence, we believe the complexity of this group is relatively medium. The third group with low complexity not only had the attributes of both the previous groups, but also received personalized product ranking based on the stated preferences of the user. This reduced the uncertainty in the information as the products were personalized by mapping the attribute preferences of the user with the average attribute-level sentiment scores from the review and hence, we believe the complexity of this group is relatively low.

The experimental part of our study has been broken down into two parts – first, text mining of reviews to extract product attributes and computing of sentiment summary scores; second, a controlled experiment to evaluate the impact of different types of information on consumers' decision process.

6.1 Extracting Product Attributes and Sentiment Scoring

Consumers in general hold a belief about a product before visiting an e-commerce platform. The belief follows a normal distribution and the belief distribution could get updated to a different mean as the consumer receives new information (Archak et al. 2011) and this new information can be attribute specific information.

Many techniques are available for extracting product attributes from product reviews. We use part-of-speech tagging as it is widely used in prior research for attribute extraction (Archak et al. 2011).

We use a Hidden Markov Model (HMM) based Part-of speech (POS) tagger. The HMM POS tagger used for feature extraction is consistent with the one discussed in prior literature (Archak et al. 2011). After extracting features, each review was tagged using a bag-of-words approach (Blair-Goldensohn et al. 2008). The sentiment score was then computed using the publicly available pre-trained library released by Google (Namrata et al. 2007). The pre-trained models are available in opensource (Python) and we chose that because, it gives both a polarity score (−1 to +1 scale) and a subjectivity score (0 to 1 scale). The subjectivity score helps to remove fake reviews as they tend to have very high subjective score, tending to be flattery.

We used more than 12,000 reviews spread across 40 different products comprising of 28 mobile phones and 12 tablets available from a large Indian e-commerce company. We chose these products as they have clearly identifiable attributes and are familiar to college students, who participated in our experiment. After applying part of speech tagging to the reviews, the nouns and noun phrases were identified from the reviews. Top 20% of the nouns and noun phrases were taken as the product attributes of the two products. We also added manufacturer defined attributes to the set of attributes extracted from POS tagging. Thus, our attribute set was a combination of both user defined and manufacturer defined attributes. In the case of cell phone, manufacturer defined attributes (which also occurred in the attributes extracted from reviews) included camera, battery and display, whereas user defined attributes included price (value for money), quality and music. The final set of 11 attributes includes battery, device, display, internet, music, os (operating system), price (value for money), camera, screen, quality and others. A bag-of-words approach was then used which had these extracted features as reference to tag every review for the list of attributes contained in the reviews.

After the tagging, a sentiment score was computed for attributes in the review. Here there is a challenge that one review can opine about more than one attribute. To overcome this challenge, for reviews that had more than one attribute tag, a phrase was formed by joining three words to the right of the attribute word and three words to the left of the attribute word (Narayanan et al. 2009). For example, as seen in Table 1, the bag-of-words approach used in this study tags the review with the attributes battery, camera and price. The tagged associated words are battery, camera, pictures, price and budget. The algorithm identifies pictures as an associated word of camera and budget as an associated word of price. As discussed earlier, a phrase is created by joining 3 words before and after the tagged word and then this phrase is used for further sentiment and subjectivity computation using the algorithm, as in prior work (Godbole et al., 2007). The results can also be seen in Table 1 shown below (Figs. 1 and 2).

A score was computed for these phrases formed, which gave the sentiment score for that particular attribute. The sentiment scores ranged from −1 to +1. The average of the sentiment scores for a particular attribute would be the summary sentiment score for that particular attribute. In this way, average sentiment score was calculated for all the extracted product attributes. A summary chart comprising of data bars for average sentiment scores for each attribute was displayed along with the product image for respondents in the medium and low complexity groups.

In the example given below, although the word camera is not used in the first review and price not used in the second review, the algorithm correctly tags the respective reviews as battery and price. The sentiment scores can also be seen in the example.

> "very worst, charger suddenly decrease when am not using any work on my mobile"
> Sentiment Score -0.55, Feature Tag – battery
>
> "excellent phone in this budget . this phone overtakes every phone in this budget range. a good comeback by acer . keep going acer... no words to say"
> Sentiment Score +0.85, Feature Tag – price

6.2 Personalization Using Review Summaries

We develop a novel attribute level personalization approach, namely, namely, personalized review feedback, which extends web personalization to online product reviews. The previous section provided an approach to generate summarized reviews in the form of a visual display using POS tagging with sentiment scores for every attribute of the product, and subsequently this section uses conjoint analysis to build the proposed personalization approach, which to the best of our knowledge, has not been explored in the domain of web personalization.

Further, in order to have a personalization strategy based on most relevant attributes, we selected top 5% of the attributes from the reviews in terms of frequency which resulted in three attributes (or factors) - battery, camera and price. Each of these three factors had two levels – high sentiment score and low sentiment score. These levels describe the sentiment score of the respective attributes which are proxies for opinions coming from product reviews for the respective attributes.

Table 1 Example of window-size based sentiment extraction

S.No	Sentence	Tag	Sentiment (−1 to + 1)	Subjectivity (0 to 1)
1	*"has an excellent battery life and very"*	battery	0.69	0.66
2	*"and very good camera that clicks amazing"*	camera	0.76	0.84
3	*"that clicks amazing pictures if you are"*	camera	0.66	0.80
4	*"pictures but the price is so costly"*	price	−0.24	0.58
5	*"costly not a budget phone!"*	price	−0.11	0.20

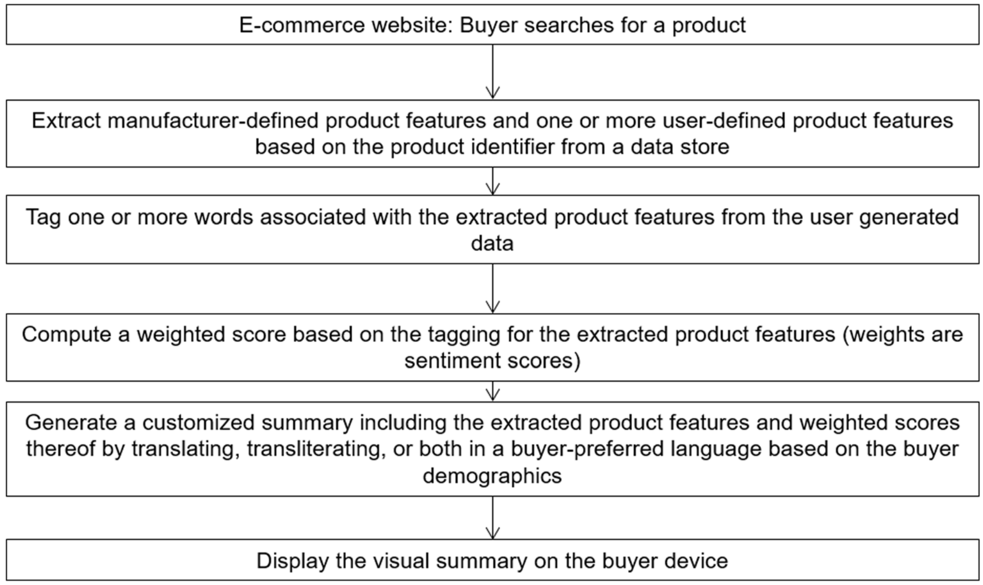


Fig. 1 Extraction of product attributes (features)

We use conjoint analysis for personalization of review summaries. This personalization is based on the attribute level average sentiment scores of the products and conjoint analysis converted the consumer stimuli to utility scores. Conjoint analysis is widely used in marketing research and also in other domains to analyze consumers' stated preferences (Green and Srinivasan 1978). There are two ways to elicit consumer preferences – stated preferences and revealed preferences. Stated preferences are the preferences explicitly stated by the consumer. Revealed preferences are those preferences that are extracted from the consumer without one's knowledge, through browsing history, purchase history etc. In this study, we apply conjoint analysis on stated preferences to arrive at a utility score for each product. Here the participants ranked 8 combinations of the product attributes with their levels in the order of their preferences. Utility scores were calculated using conjoint analysis for each product which were specific to individual consumers based on their stated preferences.

As seen in the Fig. 3 above, the top 5% of the frequently occurring attributes retrieved (3 attributes: battery, camera and price) are chosen for extracting stated preferences from the user. Each attribute has two levels – positive and negative; and there are 8 ($2^3 = 8$) different products which are ranked by the end user. These ranks are later used to calculate the part-worth utility scores for each user for every product.

The products were ranked in the order of stated consumer preferences of attributes from review summaries. The ranking was done in descending order of utility scores, thus

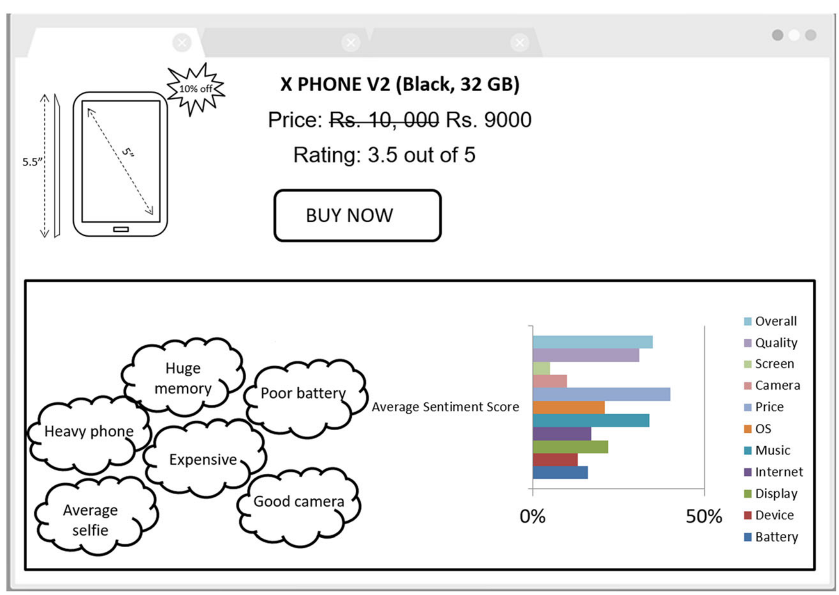


Fig. 2 Visualization of review summaries (Example: Mobile Phone)

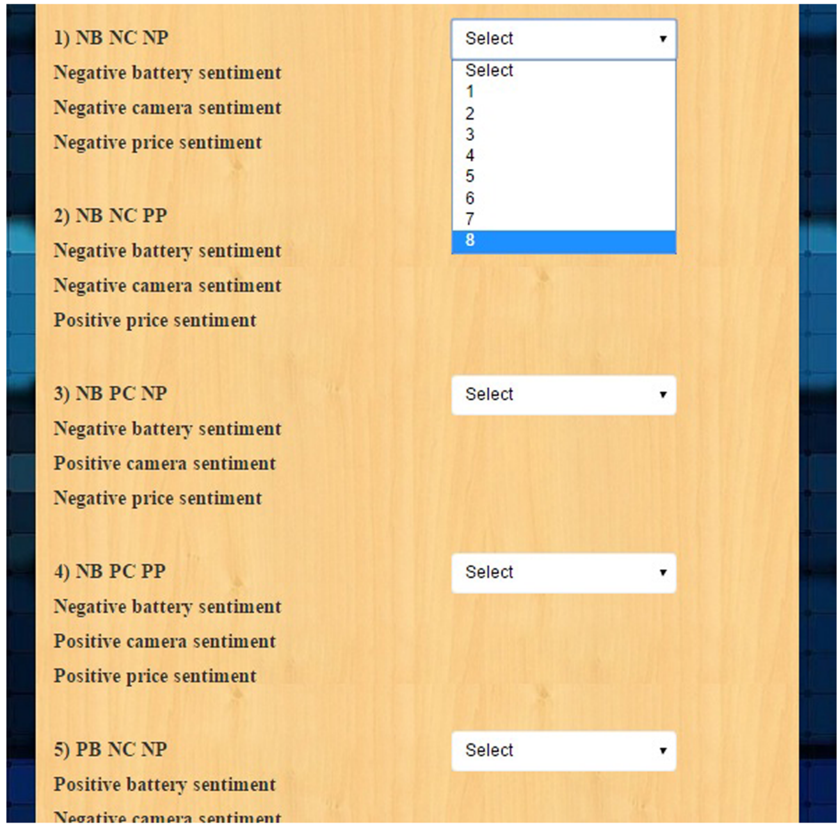


Fig. 3 Example of extracting users' attribute preferences

personalizing the webpage. By listing products in the order of consumer's preference, we formed content relevance in personalization (Tam and Ho 2006) for the group with low complexity.

6.3 Experimental Procedure

Three different versions of a virtual e-commerce website were developed for conducting the experiment to create three different levels of complexity: high complexity with voluminous raw textual reviews, medium complexity with attribute-level review summaries in the form of sentiment scores and low complexity with personalized product ranking (personalization based on reviews). The web page with unstructured reviews displayed raw product reviews in textual form, structured reviews showed aspect level summarization of reviews and personalized product ranking does personalization based on review summaries to the user. All the three websites had similar user interface design and same 40 products comprising 28 mobiles and 12 tablets. Further, the three websites also provided textual reviews with helpfulness score, product ratings, number of votes and manufacturer defined product description of the products. Participants had to sign up, create an account and also fill a form for basic demographic data. In the personalized setting, they needed to rank 8 attribute-level combinations for the products (Appendix 2).

The experiment was conducted in the IT Lab of a premier Indian institute for higher education in technology. The sample consisted of graduate students from the Institute who were willing to participate in the experiment. The procedure followed a randomized controlled experiment. The students were randomly assigned to groups using lottery method, by asking them to pick numbers randomly from a lot. Before the start of the experiment, their working memory was cleared by playing fun videos and having other entertaining activities for some time. It was made sure that all the participants started the experiment at the same time. The students were given a goal of buying 2 products mandatorily which could be mobiles or tablets. After the experiment, they were asked to checkout. The participants were given cookies and cold drinks following the experiment. There was also a lucky draw and winners were given coupons that could be redeemed.

6.4 Experiment Subjects

The experiment was conducted in multiple batches and we ensured homogeneity of the participants assigned across different batches. Each batch consisted of around 25–30 participants and the participants were equally assigned in random to the different groups. The experiment was conducted in a premier technology institute in India and subjects were graduate and research students. The eligibility to take part in the experiment was to have 5 products bought online in the past. The participants were comparable in terms of intellectual capability and online shopping experience. Research students who were greater than 40 years of age were allocated equally to different batches in a way to ensure equal split among age differences in multiple batches. We had participants aged between 23 and 56 years and an average age of 28 years. The participants were also split equally between gender in allocation to batches. There were multiple incentives for the participants to take part in the experiment.

6.5 Click Stream Data

We recorded the clickstream data in a back-end MySQL database and used them as measures of consumers' decision process (Tam and Ho 2006). In order to evaluate the uncertainty experienced by the users during the decision-making process, we measured "cart-updates" for every user during the experiment. Cart-updates is the number of changes made by the user in his or her product cart before making the buying decision.

In order to measure the effectiveness of the user's decision process in terms of acquiring relevant information and the efficiency of the user's decision process in terms of not losing time on irrelevant information, we draw upon the cognitive stopping rule. We recorded the total number of clicks made by the user during the experiment which formed a measure of the information seeking behavior of the user (Tam and Ho 2006). We also recorded the total time spent by the user to make her or his buying decision, following the approach used by Tam and Ho (2006). The total number of clicks indicate the information seeking behavior and hence measures effectiveness in terms of acquiring *relevant* information; whereas the total time spent by the user before making the buying decision indicates the efficiency in terms of not wasting time on *irrelevant* information.

6.6 Control Measures

Several control measures were adopted during the design of our experiment. Among the three different websites, the website with raw textual reviews (high complexity) was the control group and the other two websites (medium, low complexity) were the treatment group. There were several control measures that were built into the experiment.

6.6.1 Controls in Experimental Design

The three different websites were controlled for aesthetics, layout, user experience and functionalities like product search, price filter, and brand filter, etc. We controlled for the products and product details for all the groups by displaying the same products, price and product details. The scenario, sign-up process and the checkout process were also the same for all the groups. The product portfolio of the websites did not have the top 10 best-selling mobile phones to eliminate affinity towards best-selling products (this is based on the sales rank of top 10 products in two of India's biggest online retailers).

The webpages were designed such that they matched the look and feel of a typical Indian e-commerce website, and they also matched the basic structure followed in prior work involving e-commerce experiments (Balan and Mathew 2016b; Ho and Bodoff 2014; Jiang and Benbasat 2007; Liu and Karahanna, 2017; Tam and Ho, 2005, 2006). We provided all the features of an e-commerce website like product search box, brand filter, price filter, cart, checkout process, etc. In the pilot test, we also asked for a feedback about the website and incorporated the changes to imitate a real website to the best extent possible.

6.6.2 Controls in Experimental Procedure

The experimental procedure we followed had several control measures to randomize group assignment, order effect and demand effect. We followed a between-subjects experiment procedure so that the users exposed to each of these groups are different. The experiment subjects were assigned to different group using the lottery-method. The volunteers who helped us conduct the experiment had been through training sessions to explain the experiment procedure and to not cue the users on the context of the study, thus controlling for demand effect.

7 Data Analyses

Our sample consisted of 272 graduate students and after cleaning the data by dropping incomplete data and outliers, 237 cases were left for further analysis. The task complexity involved and the effectiveness and efficiency of buying decision process were measured by three variables - clicks, cart updates and time.

The sample size in all the three groups were almost equal (87, 80, 70) and to test for the correlations among the three dependent variables clicks, updates and time, we conducted Bartlett's test of sphericity. The test reported a significant degree of inter-correlation ($p < 0.001$) between the dependent variables. We tested our hypotheses using MANOVA with the clicks, cart updates and time as dependent variables, and information complexity (with respect to review information) as the independent variable having three levels. The MANOVA test showed a significant multivariate main effect for information complexity: Wilks' $\lambda = 0.906$, F (6, 464.00) = 3.910, $p = 0.001$, $\eta^2 = 0.048$, power to detect the effect = 0.969. This result shows that information complexity has a significant effect on clicks, cart updates and time.

Given the significance of the overall test, the univariate main effects were examined subsequently. Significant subject effects were found between task complexity and clicks which measures effectiveness of the consumer's decision process, $F(2, 234) = 4.118$, $p = 0.017$, partial eta square = 0.034, power to estimate the effect = 0.725; between task complexity and cart updates which measures uncertainty (in terms of difference between information required and information gathered), $F(2, 234) = 4.966$, $p = 0.008$, partial eta square = 0.041, power to estimate the effect = 0.807; and between task complexity and time which measures efficiency of the consumer (in terms of not wasting time on unnecessary information), $F(2, 234) = 6.444$, $p = 0.002$, partial eta square = .052, power to estimate the effect = 0.902.

7.1 Uncertainty (Cart Updates)

We conducted post-hoc Tukey HSD test to examine the effect of different levels of task complexity (high, medium and low) on information uncertainty. The test reported significant mean differences between raw textual reviews having high complexity (M = 2.10, SD = 1.035) and summarized reviews having medium complexity (M = 2.66, SD = 1.987), $p = 0.028$. Here again H1a relationship is significant but it is counter directional (the results are shown in detail in Table 2). This indicates that the consumers who got summarized reviews made more number of changes in the product cart before making the buying decision than the consumers who were given textual reviews.

The post-hoc Tukey HSD test also reported significant mean difference between the group that just received summarized review summaries having medium complexity (M = 2.66, SD = 1.987) and the group that received personalized product listing based on reviews having low complexity (M = 2.01, SD = 0.860), $p = 0.014$. Thus, hypothesis H1b was supported, which posits that consumers who get personalized product ranking based on reviews make fewer changes in the carts than consumers who get just summarized reviews (the results are shown in detail in Table 2).

7.2 Effectiveness of Online Buying Decision Process (Clicks)

The differences in clicks among the three groups with three different levels of task complexity were evaluated in post-hoc comparisons using Tukey HSD test. The test reported significant mean differences for clicks between high complexity (M = 3.28, SD = 4.677), and medium complexity (M = 5.04, SD = 6.103), $p = 0.049$, which is between consumers who received raw textual reviews (high complexity) and consumers who were exposed to summarized reviews (medium complexity). For hypothesis H3a relationship this result showed significant difference between the two groups, the relationship was counter directional (the results are shown in detail in Table 3). The findings suggested that consumers who were exposed to summarized reviews which were summarized clicked more than the consumers who were given textual reviews which were more complex.

The post-hoc Tukey HSD test also showed significant difference between the group having medium complexity with summarized reviews (M = 5.04, SD = 6.103) and the group having low complexity with personalized product ranking based on reviews (M = 3.01, SD = 2.856), $p = 0.028$. This result supported hypothesis H3b, which posited that consumers who get personalized product ranking based on reviews seek lesser information than consumers who get summarized reviews with no personalization (the results are shown in detail in Table 3).

7.3 Efficiency of Online Buying Decision Process (Time)

Time measures the overall time spent by the consumer in the e-commerce platform, from the time one logged into the e-commerce web page to the time one checked out the products. The post-hoc Tukey HSD test reported significant mean difference between the group that received raw textual reviews having high complexity (M = 358.39, SD = 207.03) and the group that received summarized reviews having medium complexity (M = 438.03, SD = 186.21), $p = 0.016$. This result shows H2a as significant but it is counter directional again (the results are shown in detail in Table 4).

The post-hoc Tukey HSD test also showed significant difference between the group that received summarized reviews with medium complexity (M = 438.03, SD = 186.21) and the group that had personalization based on review summaries having low complexity (M = 337.01, SD = 149.70), $p = 0.003$. This result supported our hypothesis H2b, which posits

Table 2 Tukey HSD test with cart updates as dependent variable

Dependent variable – Cart Updates (Uncertainty in decision process)

Complexity		Mean	Standard Deviation	Mean Difference	Sig.	Remarks
High vs Medium	High Complexity Medium Complexity	2.10 2.66	1.035 1.987	- 0.56	0.028	H1a significant, but counter directional
Medium vs Low	Medium Complexity Low Complexity	2.66 2.01	1.987 0.860	0.65	0.014	H1b supported

Table 3 Tukey HSD test with clicks as dependent variable

Dependent variable – Clicks (Effectiveness of the decision process)						
Complexity		Mean	Standard Deviation	Mean Difference	Sig.	Remarks
High vs Medium	High Complexity Medium Complexity	3.28 5.04	207.032 186.210	−1.76	0.049	H3a significant, but counter directional
Medium vs Low	Medium Complexity Low Complexity	5.04 3.01	186.210 149.698	2.03	0.028	H3b supported

that consumers who get personalized product ranking based on reviews spend lesser time to make their buying decision than consumers who get review summaries with no personalization (the results are shown in detail in Table 4).

8 Discussion

Our study seeks to understand the effect of complex information environment on online consumers' decision process reflected in information seeking behavior, effectiveness (in terms of acquiring accurate information) and efficiency (in terms of not wasting time on unnecessary information). Two pair-wise comparisons: high vs medium complexity and medium vs low complexity are done for uncertainty (cart changes: H1a,b), efficiency (time: H2a,b) and effectiveness (clicks: H3a,b). This study hypothesizes the uncertainty, efficiency and effectiveness of decision-making process will increase from high-low complexity with personalized review feedback group being the most effective and efficient. Three (H1b, H2b, H3b) out of six hypotheses are fully supported showing that personalized product ranking based on review summaries enable consumers in making choice decisions effectively and efficiently. Contrary to our expectations, H1a, H2a and H3a are not supported in the same direction they were hypothesized although the results were significant. These findings suggest that the groups which have review summaries (medium complexity) spend more time to make their buying decision than the group exposed to raw textual reviews (high complexity). We further conducted a focus group discussion with participants from the lab experiment to investigate and explain the counter intuitive results.

8.1 Focus Group Discussion

In order to explore the thought processes involved in the consumers' decision making and to understand the consumer behavior, a Focus Group Discussion (FGD) was conducted with participants after the experiment. The FGD was conducted in batches of 8–10 people in three rounds. We used a semi-structured questionnaire to guide the FGD process. Various problems and experiences shared by the consumer were analyzed and further compared with extant literature.

8.1.1 Textual Reviews and Summarized Reviews

Although participants who received raw textual reviews required more cognitive efforts to process voluminous information and make a buying decision than participants who received review summaries, our results show that the former clicked less number of times, added less number of products to the cart and took less time in decision making. In seeking explanation for this apparently counterintuitive finding we noted from our FGDs that the participants did not spend time reading the textual reviews as it was voluminous. They have too much of information available in the webpages leading to information overload and difficulty in extracting relevant information. This response is consistent with the findings of Payne (1976) who reported that "the percentage of information searched declined both as the number of alternatives in a decision situation increased and as the number of dimensions per alternative increased." The working memory of the human brain has limited capacity and the learning ability deteriorates as information overloads (Baddeley 1992). Furthermore, if the amount of information processing exceeds a certain limit, the demand for cognitive

Table 4 Tukey HSD test with time as dependent variable

Dependent variable – Time (Efficiency of the decision process)						
Complexity		Mean	Standard Deviation	Mean Difference	Sig.	Remarks
High vs Medium	High Complexity Medium Complexity	358.39 438.03	207.032 186.210	- 79.64	0.016	H2a significant, but counter directional
Medium vs Low	Medium Complexity Low Complexity	438.03 337.01	186.210 149.698	101.02	0.003	H2b supported

effort may force people to simplify their task execution strategies or they would even sacrifice their task performance for reducing effort (Jiang and Benbasat 2007; Kahneman 1973b). These findings from literature along with the observations emerging out of our FGD data explain the significantly different lower cognitive effort indicated through mean value of clicks, cart updates and time for the raw textual reviews group as compared to the summarized reviews group.

We also examined the decision model adopted by the participants in our experiment. Participants in the FGD revealed that the review summarization helped them to make decisions quickly. According to Payne (pp. 367, Payne 1976) consumers adopt different types of decision models to make their buying decision. Some consumers adopt a compensatory model where the consumer arrives at a decision by evaluating all the attributes of a product and choosing the one with most positive and least negative attributes. In contrast, some others adopt a non-compensatory model where the consumers will not buy a product till they encounter one with the attributes they want. In non-compensatory model, the consumers assume that the product is not worth the money if it does not have the attributes they need. The conjunctive model and elimination-by-aspect model are few of the prominent non-compensatory decision models. Conjunctive model involves setting up of minimum values for every attribute in the product and if the consumer evaluation exceeds that, buying decision will be made by the consumer. In elimination-by-aspect model, the decision maker processes one cue at a time and gradually reduces the number of alternatives until only one option is left. All these decision models require the consumers to compare different products before the buying decision (Payne 1976). Our FGD data showed that in the summarized review treatment, consumers tend to click more products and spend more time comparing different products. They find relevant cues of the product from the summary chart and now they use these cues to find the best product by comparing it with others.

8.1.2 Effectiveness and Efficiency in Online Buying Decision Process

The effectiveness of the consumer (in terms of acquiring accurate information) and efficiency of the consumer (in terms of not wasting time on unnecessary information) is the highest in the group with personalized review feedback. This group is the most effective as it required the least number of clicks to make the buying decision (M = 3.01), the most efficient as it required the least time to make the buying decision (M = 337.01) and the group also has lowest complexity associated with the buying task as it has the least information uncertainty indicated by the least number of changes in the product cart before making the buying decision (M = 2.01). However, providing with a personalized review feedback could potentially reduce the opportunity to explore and search more products in the online environment. According to (Fong 2016) personalization leads to shrinking of consumer search resulting in reduced cross-selling.

The group with summarized reviews has the highest clicks (M = 5.04,), highest time (M = 438.02) and highest number of changed in product cart before making the buying decision (M = 2.66). Although this group has least effectiveness, least efficiency and highest information uncertainty, the FGD data revealed that consumers found the summarized reviews to be very helpful and the new information provided by the review summaries in the form of sentiment scores motivated the users to compare multiple products with this information and this resulted in increased user activity in this group thus leading them to spend more time to make their buying decision. Thus, though this group required more efforts from the user in terms of clicks and time, the FGD revealed that this group enabled the consumers to compare multiple products and consumers in this group had the opportunity to explore more products.

8.1.3 Counter Intuitive Findings

Drawing upon Browne et al. (2007), we hypothesized that the stopping rules invoked by the users to terminate information search will be contingent upon the acquisition of relevant information. We argued that the complexity of information should reduce as we make the information acquisition easier, as visual and tabular presentation are known to bring better comprehension and lesser task complexity (Jiang and Benbasat 2007). Our results show that the decisions models proposed for consumer choices proposed to form hypotheses in this study were not suitable for personalized environment, as explained in Section 8.1. The results of this study are also complementary to recent findings in information systems literature on influence of the online environment on users' tradeoff between item sampling and item selection (Ho and Bodoff 2014); and users' preferences (Liu and Karahanna 2017).

Although we get counter-intuitive findings from the data analyses, the focus group discussion reveals that users find easier to comprehend relevant information when presented with summaries and it is their curiosity to use those scores to compare different choice sets that had led to the increase in the number of clicks, time and cart-updates. This is another complementary finding from this work, which provides a new direction of future research which should could focus upon the optimal choice sets in the E-commerce environment.

9 Theoretical Implications

This paper makes a few important contributions to the literature on web personalization. The main contribution of this study is in the synthesis of a theory for web personalization based on the theory of complexity and cognitive stopping rule and the testing of the proposed theory using experimental data. We developed evidence to suggest that review-based web personalization significantly impacts consumer decision process.

We also made certain important methodological contributions in conducting this study. First, a novel attribute level personalization approach, namely, personalized review feedback, which extends web personalization to online product reviews has been developed. Second, the study empirically evaluates the impact of attribute-level review summaries on the buying behavior of a consumer. Third, it provides an approach to generate structured reviews in the form of visual summary with sentiment scores for every attribute of a product. Fourth, we use conjoint analysis to implement reviews-based personalization, which has not been explored in the domain of personalization to our knowledge. The early findings from this study can be extended to the context of different types of goods and different personalization approaches for future research.

10 Managerial Implications

Our findings have important implications for recommender system strategies and personalization in e-commerce. The results of this study generated empirical evidence to show that consumers decision process become effective and efficient when they are provided with a personalized review feedback, which is personalized product listing based on review summaries. The e-commerce platforms can provide both review summaries using sentiment scores and also personalize the product listing by positioning the products that have a good review feedback for consumers' preferred attributes in the top. By doing so, the search cost for the consumers become much lesser. Personalization strategy of this kind would be best suited for *hurried shoppers* who want to close the shopping activity quickly. However, opening up personalization for all categories of customers could negatively impact the potential for cross selling.

Another key finding from this study is that consumers exhibit higher levels of online activity when presented with summarized reviews in the form of sentiment scores. Thus, summarized reviews will be a great opportunity for e-commerce platforms to stimulate online activity. This could also be of great use to address the cold-start problem in e-commerce. The consumers get an opportunity to explore more products when presented with review summaries, as they are motivated to compare many products using the summarized information. Here effectiveness and efficiency of the consumer decision process are inversely related to the opportunity to explore new products.

In sum, if the e-commerce platform wants to promote cross-selling or want to address the cold-start problem, they can provide summarized reviews with sentiment scores. This will stimulate more online activity and consumers will get the opportunity to explore new products as they start to click and compare more products before they make the buying decision. If the e-commerce platform wants to increase the effectiveness and efficiency of the consumers' buying decision process, the platform can provide a personalized review feedback, which uses the consumers' preferences and personalizes the product listing by positioning products that have higher sentiment scores for the consumers' preferred attributes.

11 Future Scope for Research

One key insight from this work is that although effectiveness (number of clicks) and efficiency (time spent) of the decision process was the lowest when consumers were exposed to personalized review feedback, the metrics are not much different between no personalization group and personalized review feedback group. This pattern was also consistent with the third dependent variable: uncertainty in information acquisition (number of changes in product cart) in both these groups.

The FGD revealed that users did not read much when subject to information overload and hence made purchase decision using some rule-of-thumb or heuristic. In contrast, when provided with review summary, the users were motivated to search for the product that is the most suitable for them. In the personalized review feedback environment, this search cost is reduced as products were personalized according to the attribute preferences of the user. This forms an inverted U-shaped curve, consistent with the findings of Payne (1976) and Kahneman (1973c).

Future research can extend these findings to study how consistent these results hold in different categories of products. The current research uses attribute rich products (mobile phone, tablets) which are search goods; it would also be interesting to study this for experience and service goods.

12 Conclusion

Our study analyzes the impact of summarized reviews and personalized review feedback on consumers' buying decision process. Previous studies have evaluated the impact of personalization and online word of mouth on sales, but the integration of both remains unexplored in IS research. Our work integrates concepts from information systems, computer science (more precisely NLP for POS tagging and sentiment scoring) and social-cognition to study the impact of content relevance and complexity on consumers' decision process. The results of this study show that consumers have high effectiveness (in terms of acquiring accurate information) and efficiency (in terms of not wasting time on unnecessary information) in making their buying decisions in our proposed personalized review feedback-based e-commerce environment. On the other hand, consumers who receive summarized reviews in the form of attribute-level sentiment scores are motivated to put more efforts to seek relevant information to make choices based on their preferences. This is reflected in the additional effort they expense through click stream data and FGDs. Finally, the groups who receive raw textual reviews avoid higher cognitive efforts and resort to decision making without information, which leads to inferior decision quality.

Our work is restricted to products like mobile phones and tablets which may not be comparable to all product categories. This being a limitation of this study, separate investigation is imperative for other product categories such as commodities, experience goods and services. The experiment required purchase goals for all participants and behavior in the absence of a purchase goal is not evaluated. The consumption of review information by a consumer who is browsing through the webpages without any purchase goal could be different from the consumer who searches with specific goals. Thus, goal specificity could be further introduced as an additional factor in the experiment. Furthermore, the type of personalization used in an e-commerce website also affects consumer behavior. If the webpage's personalization remains the same over time, it is static personalization. In contrast to it, if the personalization improves over time by learning the user more, it is dynamic or adaptive personalization. Our experiment used static personalization using conjoint analysis and the impact in a dynamically personalized environment requires further investigation. The time of introduction of personalization based on reviews could also influence a buyer's decision process. More the time taken to introduce personalization, the accuracy of personalization increases but the probability of the consumer buying the recommended product could decrease. The best time to introduce personalization in online consumers' decision process could be further addressed in future research.

Appendix 1 - Stopping Rule

Ho and Bodoff (2014) adapted DeGroot's (1968) model to determine stopping rule of consumers in the online context. The model follows a Bayesian approach and considers the fixed cost spent by the user to sample more items and also the subjective value given to the item by the user. Consumers are said to have some beliefs about the product in their working memory while visiting the e-commerce platform to make their purchase (Archak et al. 2011). These beliefs follow a normal distribution with μ as the actual mean value (subjective) of the item and τ as the variance; and after receiving the information available in the website, these beliefs get updated to newer mean values $\mu 0$ and $\tau 0$. Finally, they invest some cognitive efforts to process these beliefs in their working memory and make a buying decision. Let us denote θ as $\mathbf{N(\mu, \tau)}$ where μ is the mean value (subjective) of the item and $\mathbf{x}$ is the mean value of the additional item sampled. Now, the updated estimate $\theta^{\wedge}$ is $\mathbf{N([(\tau\mu + x)/(\tau + r)], \tau + r)}$, where r is precision (Ho and Bodoff 2014). If v is the best estimated mean value of the item, then according to the Bayesian stopping rule a consumer will sample items until

$$v > \mu + f\left(\frac{\tau}{\tau + r}, c\right)$$

Thus, in a personalized environment, the fixed cost of search remains the same and the value derived from sampling additional items is less because $f\left(\frac{\tau}{\tau+r}, c\right)$ is very less.

Appendix 2 - Website Screenshots

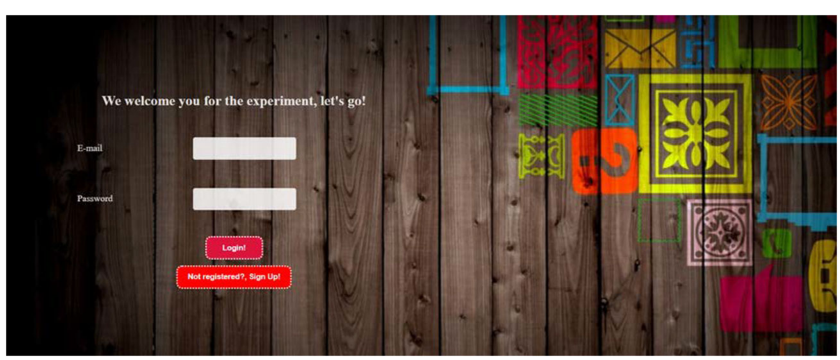


Fig. 4 Website login page for the three groups

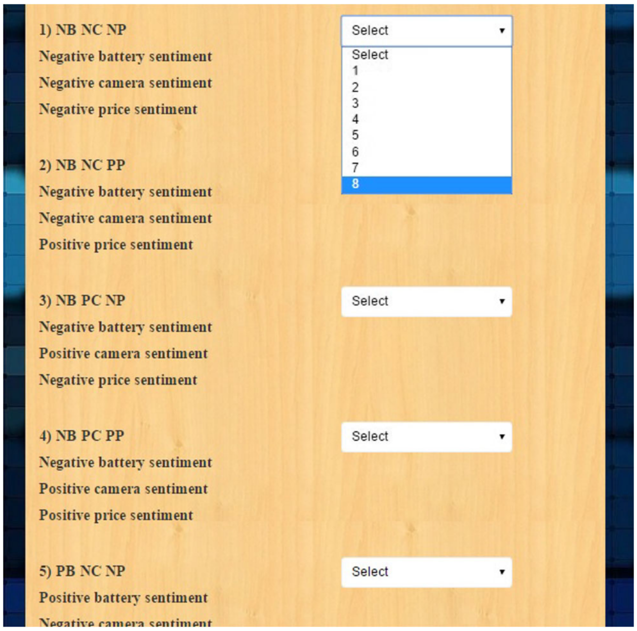


Fig. 5 Customer stimuli extracted to personalize using conjoint analysis

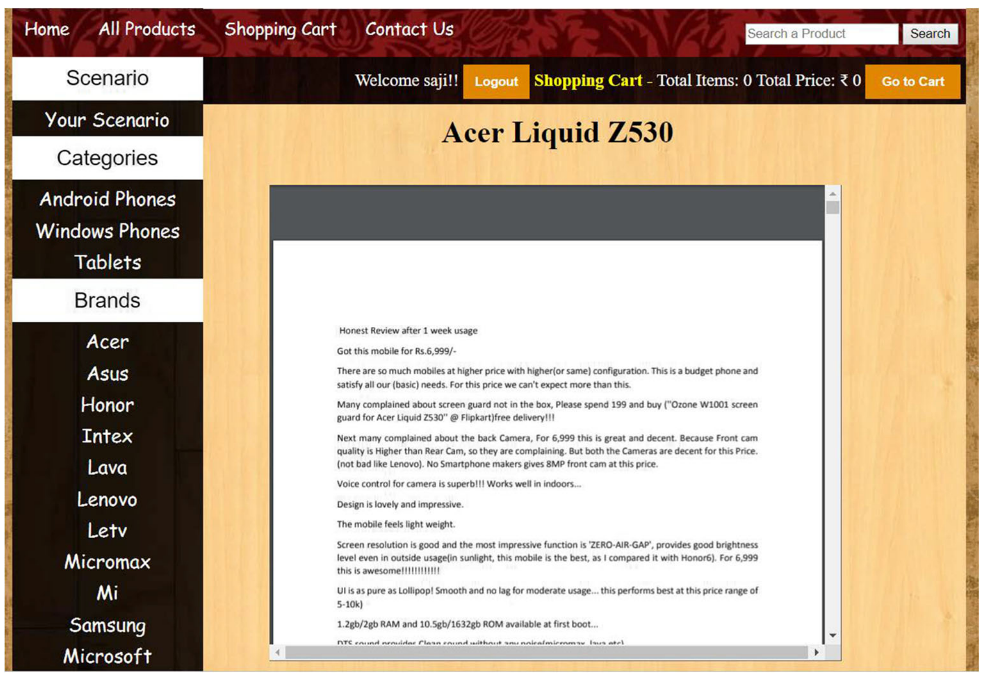


Fig. 6 Unstructured reviews group with voluminous product reviews (provided by a major online retailer in India)

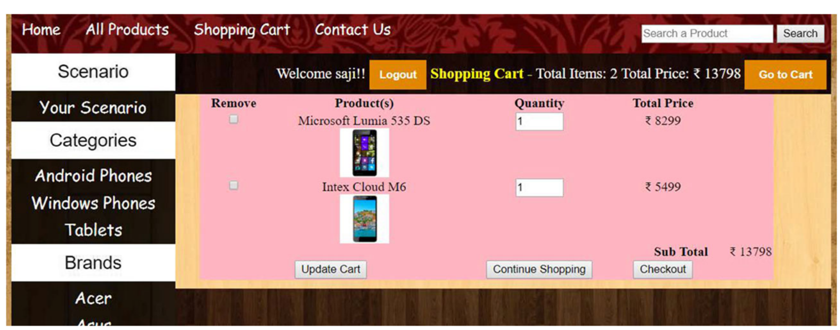
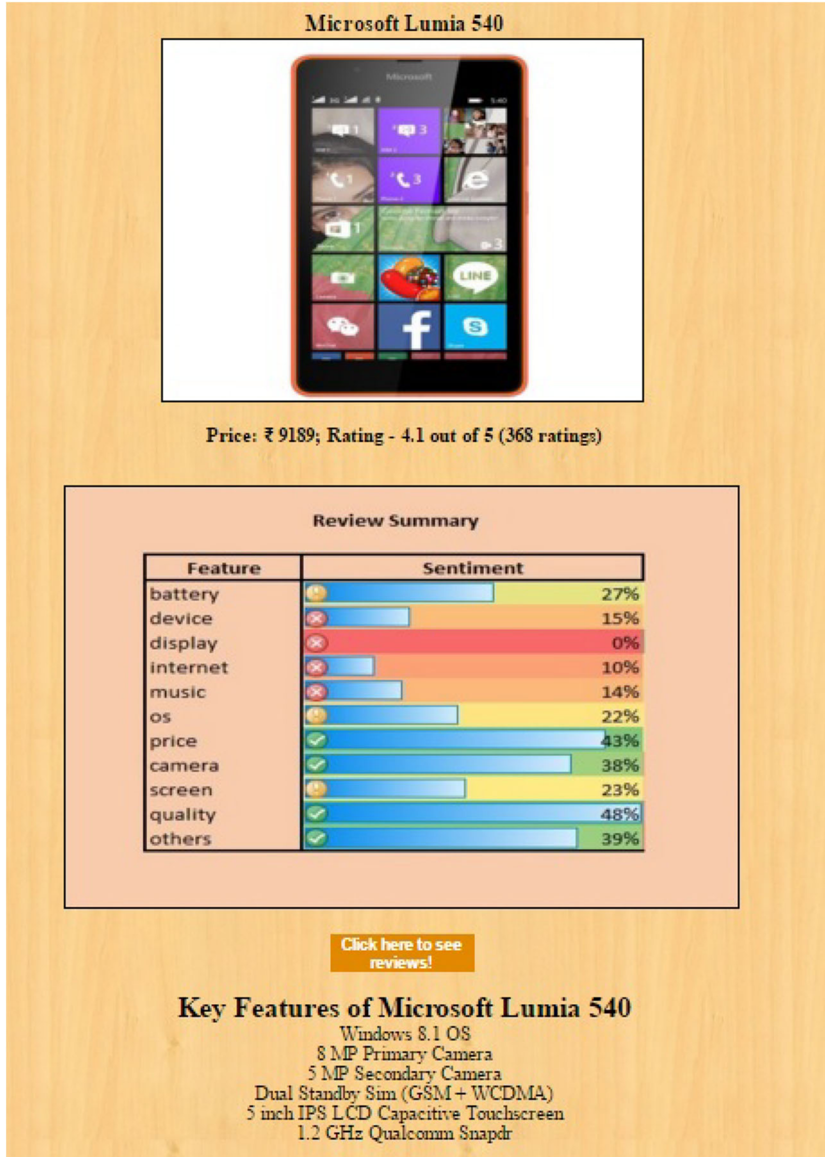


Fig. 7 Product car with update and checkout options

Fig. 8 Review summary in structured reviews group



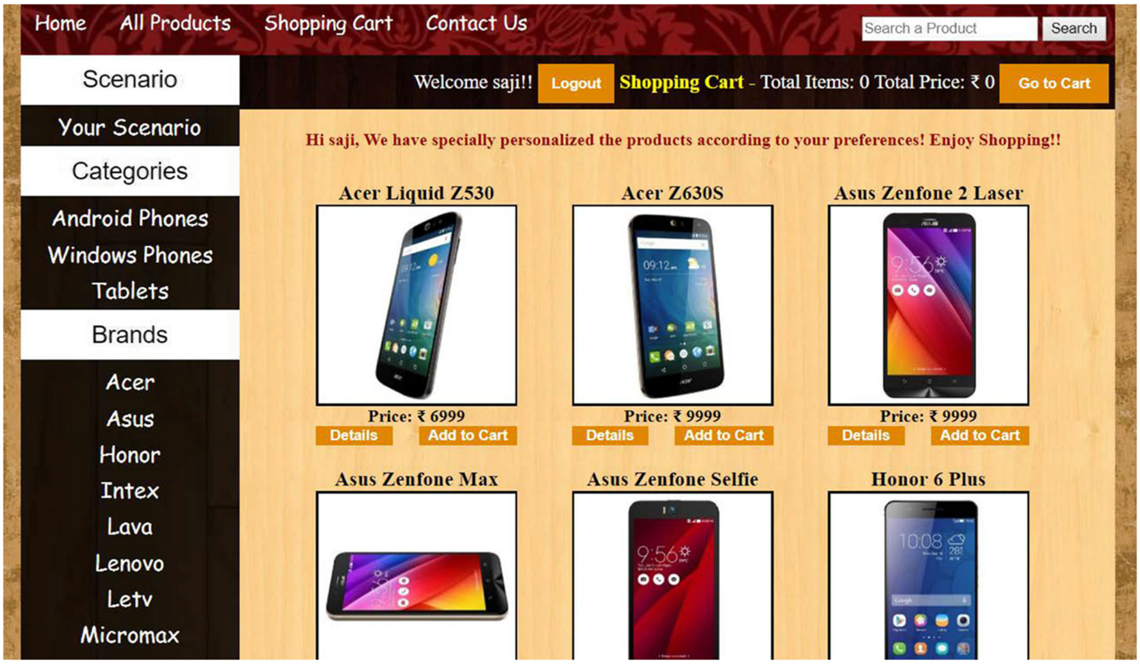


Fig. 9 Personalized product ranking in personalization based on reviews group

References

Alper, B., Yang, H., Haber, E., & Kandogan, E. (2011) Opinionblocks: Visualizing consumer reviews. In *IEEE VisWeek Workshop on Interactive Visual Text Analytics for Decision Making, Providence, USA, 2011*: IEEE.

Archak, N., Ghose, A., & Ipeirotis, P. G. (2011). Deriving the pricing power of product attributes by mining consumer reviews. *Management Science, 57*(8), 1485–1509.

Arndt, J. (1967). Role of product-related conversations in the diffusion of a new product. *Journal of Marketing Research, 4*(3), 291–295.

Baddeley, A. (1992). Working memory and conscious awareness. *Theories of memory*, 11–20.

Balan, U. M., & Mathew, S. (2015) Online word of mouth using text mining: A review of literature and future directions. In *Computational Intelligence: Theories, Applications and Future Directions (WCI), 2015 IEEE Workshop on, 2015* (pp. 1–6): IEEE.

Balan, M., & Mathew, S. (2016a) Impact of personalized review summaries on buying decisions: An experimental study. In *22nd Americas Conference on Information Systems (AMCIS), San Diego, USA, 2016*: AIS.

Balan, U. M., & Mathew, S. (2016b) Impact of web personalization of online word of mouth on buyers' decision process: An experimental study. In *IEEE International Conference on Management of Innovation and Technology (ICMIT), 2016* (pp. 139–144): IEEE.

Balan, U. M., & Mathew, S. (2017) Impact of Product Review Summarization and Personalization on Online Consumers' Decisions: An Experimental Study. In *ACM SIGMIS Conference on Computers and People Research, IIM Bangalore, India, 2017*: ACM.

Beineke, P., Hastie, T., Manning, C., & Vaithyanathan, S. (2004) Exploring sentiment summarization. In *Proceedings of the AAAI spring symposium on exploring attitude and affect in text: theories and applications, 2004* (Vol. 39).

Belanger, F. (2012). Theorizing in information systems research using focus groups. *Australasian Journal of Information Systems, 17*(2), 109–135.

Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G. A., & Reynar, J. (2008) Building a sentiment summarizer for local service reviews. In *WWW workshop on NLP in the information explosion era, 2008* (Vol. 14, pp. 339–348).

Browne, G. J., & Pitts, M. G. (2004). Stopping rule use during information search in design problems. *Organizational Behavior and Human Decision Processes, 95*(2), 208–224.

Browne, G. J., Pitts, M. G., & Wetherbe, J. C. (2007). Cognitive stopping rules for terminating information search in online tasks. *MIS Quarterly*, 89–104.

Busemeyer, J. R., & Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology, 32*(2), 91–134.

Byström, K., & Järvelin, K. (1995). Task complexity affects information seeking and use. *Information Processing & Management, 31*(2), 191–213.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review, 13*(1), 40–52.

Chen, P.-Y., S.-Y. Wu and J. Yoon (2004) The impact of online recommendations and consumer feedback on sales. ICIS 2004 Proceedings, Paper 58.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345–354.

Clemons, E. K., Gao, G. G., & Hitt, L. M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of Management Information Systems, 23*(2), 149–171.

Couger, J. D. (1996). *Creativity & innovation in information systems organizations*. Danvers: Boyd & Fraser Publishing Company.

Das, A. S., Datar, M., Garg, A., & Rajaram, S. (2007) Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web, 2007* (pp. 271–280): ACM.

DeGroot, M. H. (1968). Some problems of optimal stopping. *Journal of the Royal Statistical Society: Series B (Methodological), 30*(1), 108–122.

Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing, 21*(4), 23–45.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review, 14*(4), 532–550.

Fitch, D. (2007). Structural equation modeling the use of a risk assessment instrument in child protective services. *Decision Support Systems, 42*(4), 2137–2152.

Fong, N. M. (2016). How targeting affects customer search: A field experiment. *Management Science, 63*(7), 2353–2364.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research, 19*(3), 291–313.

Frost, P. J., & Mahoney, T. A. (1976). Goal setting and the task process: I. an interactive influence on individual performance. *Organizational Behavior and Human Performance, 17*(2), 328–350.

Gao, G. G., Greenwood, B. N., Agarwal, R., & McCullough, J. S. (2015). Vocal minority and silent majority: How do online ratings reflect population perceptions of quality? *MIS Quarterly, 39*(3), 565–589.

Ghose, A., & Ipeirotis, P. G. (2007) Designing novel review ranking systems: predicting the usefulness and impact of reviews. In *Proceedings of the ninth international conference on Electronic commerce, 2007* (pp. 303–310): ACM.

Ghose, A., & Ipeirotis, P. G. (2010). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge and Data Engineering, 23*(10), 1498–1512.

Godbole, N., Srinivasaiah, M., & Skiena, S. (2007). Large-scale sentiment analysis for news and blogs. *Icwsm, 7*(21), 219–222.

Green, P. E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research, 5*(2), 103–123.

Hammond, K. R. (1986). A theoretically based review of theory and research in judgment and decision making. COLORADO UNIV AT BOULDER CENTER FOR RESEARCH ON JUDGMENT AND POLICY.

Ho, S. Y., & Bodoff, D. (2014). The effects of Web personalization on user attitude and behavior: An integration of the elaboration likelihood model and consumer search theory. *MIS Quarterly, 38*(2), 497–520.

Ho, S. Y., Bodoff, D., & Tam, K. Y. (2011). Timing of adaptive web personalization and its effects on online consumer behavior. *Information Systems Research, 22*(3), 660–679.

Hu, N., Pavlou, P. A., & Zhang, J. (2006). Can online reviews reveal a product's true quality?: Empirical findings and analytical modeling of online word-of-mouth communication. In Proceedings of the 7th ACM conference on Electronic commerce (pp. 324–330). ACM.

Hu, N., Zhang, J., & Pavlou, P. A. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM, 52*(10), 144–147.

Hu, N., Pavlou, P. A., & Zhang, J. J. (2017). On self-selection biases in online product reviews. *MIS Quarterly, 41*(2), 449–471.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology, 79*(6), 995–1006.

Jiang, Z., & Benbasat, I. (2007). The effects of presentation formats and task complexity on online consumers' product understanding. *MIS Quarterly*, 475–500.

Kahneman, D. (1973a). *Attention and effort* (Vol. 1063). Englewood Cliffs: Prentice-Hall.

Kahneman, D. (1973b). *Attention and effort*: Citeseer.

Kahneman, D. (1973c). *Attention and effort*. NJ: Prentice-Hall Englewood Cliffs.

Kim, S. M., & Hovy, E. (2004). Determining the sentiment of opinions. In Proceedings of the 20th international conference on Computational linguistics (p. 1367). Association for Computational Linguistics.

Korper, S., & Ellis, J. (2001). Thinking ahead in e-commerce. *Executive Excellence, 18*(7), 19–19.

Krishnaraju, V., Mathew, S. K., & Sugumaran, V. (2016). Web personalization for user acceptance of technology: An empirical investigation of E-government services. *Information Systems Frontiers, 18*(3), 579–595.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing, 70*(3), 74–89.

Liu, Q. B., & Karahanna, E. (2017). The dark side of reviews: The swaying effects of online product reviews on attribute Preference construction. *MIS Quarterly, 41*(2), 427–448.

Liu, J., Cao, Y., Lin, C. Y., Huang, Y., & Zhou, M. (2007). Low-quality product review detection in opinion summarization. In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL) (pp. 334–342).

Moore, D. M., Burton, J. K., & Myers, R. J. (1996). Multiple-channel communication: The theoretical and research foundations of multimedia. *Handbook of research for educational communications and technology*, 851–875.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica, 27*, 84–92.

Narayanan, R., Liu, B., & Choudhary, A. (2009) Sentiment analysis of conditional sentences. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1, 2009* (pp. 180–189): Association for Computational Linguistics.

Norman, D. A., & Bobrow, D. G. (1975). On data-limited and resource-limited processes. *Cognitive Psychology, 7*(1), 44–64.

Payne, J. W. (1976). Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organizational Behavior and Human Performance, 16*(2), 366–387.

Schroder, H. M., Driver, M. J., & Streufert, S. (1967). *Human information processing*. New York: Holt, Rinehart & Winston.

Shaw, M. E. (1971). Group dynamics: The Psychology of Small group behavior.

Sheng, H., Nah, F. F.-H., & Siau, K. (2008). An experimental study on ubiquitous commerce adoption: Impact of personalization and privacy concerns. *Journal of the Association for Information Systems, 9*(6), 344.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99–118.

Simon, H. A. (1981). *The sciences of the artificial* (rev. ed.). Cambridge: MIT Press.

Slovic, P. (1995). The construction of preference. *American Psychologist, 50*(5), 364.

Smith, G. F., Benson, P. G., & Curley, S. P. (1991). Belief, knowledge, and uncertainty: A cognitive perspective on subjective probability. *Organizational Behavior and Human Decision Processes, 48*(2), 291–321.

Stigler, G. J. (1961). The economics of information. *Journal of Political Economy, 69*(3), 213–225.

Sun, M. (2012). How does the variance of product ratings matter? *Management Science, 58*(4), 696–707.

Tam, K. Y., & Ho, S. Y. (2005). Web personalization as a persuasion strategy: An elaboration likelihood model perspective. *Information Systems Research, 16*(3), 271–291.

Tam, K. Y., & Ho, S. Y. (2006). Understanding the impact of web personalization on user information processing and decision outcomes. *MIS Quarterly, 30*4, 865–890.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review, 79*(4), 281.

Wyer Jr., R. S., & Srull, T. K. (2014). *Memory and cognition in its social context*. Hove: Psychology Press.

Yaniv, I. (2004). Receiving other people's advice: Influence and benefit. *Organizational Behavior and Human Decision Processes, 93*(1), 1–13.

Zhang, Z., & Varadarajan, B. (2006). Utility scoring of product reviews. In Proceedings of the 15th ACM international conference on Information and knowledge management (pp. 51–57). ACM.

Zhuang, L., Jing, F., & Zhu, X.-Y. (2006) Movie review mining and summarization. In *Proceedings of the 15th ACM international conference on Information and knowledge management, 2006* (pp. 43–50): ACM.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Mahesh Balan U is currently a Senior Data Scientist at ZS Associates, India. His doctoral research focuses on web personalization and consumer behavior and his other research interests include e-commerce, information privacy, behavioral economics and applied machine learning. His primary focus is on using state-of-the-art developments in computer science to solve and understand problems in information systems. He has over 4 years of industry experience in organizations like Tata Consultancy Services (TCS), Amazon and ZS Associates. He is a HICSS Doctoral Scholar. His articles have been published in reputed Information Systems journals published by ACM, Springer, AIS and IEEE. His paper titled "Impact of personalized review summaries on buying decisions: An experimental study" got the best paper award runner-up in AMCIS 2016 at San Diego, USA. He is also a very active member of AIS India Chapter (INAIS).

Saji K. Mathew is currently a Professor at the Department of Management Studies, Indian Institute of Technology Madras. His doctoral research and subsequent academic work focused on the role of Information Technology in Business and Management. As a Fulbright Scholar, he did his post-doctoral research on risk mitigation in offshore IT outsourcing at the Goizueta Business School of Emory University, Atlanta (USA). His present research interests cover information privacy and personalization strategy, digital platforms, and business value of digital technologies. His articles have been published in reputed international journals such as Information & Management, Journal of Strategic Information Systems, European Management Journal and Information Systems Frontiers. He has served as the Guest Editor for Information Systems Frontiers, and Associate Editor for ICIS-2017. He has about 10 years of work experience in the area of industrial automation in private and public-sector companies. He has also provided industrial training and consulting for companies such as Exxon Mobile, Genpact, HP Globalsoft, Oracle India, Primus Retail, L&T and Hindustan Aeronautics Limited in addition to sponsored research projects for Nissan, Hand in Hand, Infosys DSIR and DFID. He is one of the founding members of the AIS India Chapter (INAIS) and also serves as the Treasurer for INAIS.