International Journal of Information Management 39 (2018) 30–37

Contents lists available at ScienceDirect

International Journal of Information Management

journal homepage: www.elsevier.com/locate/ijinfomgt

Longer online reviews are not necessarily better

Lior Fink*, Liron Rosenfeld, Gilad Ravid

Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Israel

ARTICLE INFO

Keywords:
Online product reviews
Review length
Information overload
Mobile apps

ABSTRACT

Models of information processing have long suggested that people respond in a curvilinear manner to variation in information load and that information use may be restricted when available information is either scarce or abundant. Research on online product reviews, however, suggests that the relationship between the length of online reviews available to consumers and effectiveness measures is positive and linear. To explain this discrepancy, we argue that review length has a negative curvilinear (inverted-U-shaped) relationship with effectiveness and that such a relationship has seldom been observed in previous studies because those have analyzed data collected in low-constraint settings of information processing. The analysis of data about online reviews for free and paid apps, collected on two mobile app stores, provides consistent evidence in support of the hypothesized curvilinear relationship. The findings suggest that maximum cognitive load is experienced at lower review lengths for paid apps than for free apps and that the marginal utility for the majority of review length observations is positive or nonsignificant for free apps and negative for paid apps. These findings are consistent with product-related differences in information processing motivation. The study contributes to the ongoing debate on the ideal length of messages in electronic environments.

1. Introduction

Information processing theory suggests that information use may be restricted when the amount of information available is either scarce or abundant (Schroder, Driver, & Streufert, 1967). The latter case represents a situation of information overload, which refers to the finite limits to the ability of people to process information and to the deterioration in performance once these limits are surpassed (Gross, 1964; Jacoby, 1977; Toffler, 1970). Models of human information processing have long suggested that people respond in a typical curvilinear manner to variation in information load in their environments. As environmental load increases, the amount of information actually used in decisions increases until a point of optimum cognitive load, after which information use declines (Driver & Mock, 1975). The negative performance consequences of information overload have been confirmed in various decision-making contexts (Jacoby, Speller, & Kohn, 1974; Jones, Ravid, & Rafaeli, 2004; Malhotra, 1982; O'Reilly, 1980).

These cognitive limitations, however, are largely absent in research on the effectiveness of online consumer-generated product reviews. Driven by the motivation to understand word-of-mouth behavior in an era of electronic commerce (Ahmad & Laroche, 2017; King, Racherla, & Bush, 2014), numerous attempts have been made in the past decade to empirically capture how quantitative and qualitative characteristics of online product reviews are related to consumer behavior, in particular to product sales (e.g., Chevalier & Mayzlin, 2006; Floyd, Freling, Alhoqail, Cho, & Freling, 2014; Forman, Ghose, & Wiesenfeld, 2008; Ghose & Ipeirotis, 2011; Gu, Park, & Konana, 2012; Neirotti, Raguseo, & Paolucci, 2016). A characteristic often observed in such empirical investigations is review length, typically measured as the average number of characters or words included in reviews for a specific product. This characteristic is consistently hypothesized to be positively related to either product sales or review helpfulness (Baek, Ahn, & Choi, 2012; Fang, Zhang, Bao, & Zhu, 2013; Mudambi & Schuff, 2010). The reasoning underlying these hypotheses rests on two arguments. The first is that longer reviews are less likely to be overlooked than shorter reviews because they take up more screen space and are visually more salient (Kuan, Hui, Prasarnphanich, & Lai, 2015). The second argument is that compared to shorter reviews, longer reviews contain more product-related information, which is likely to increase the consumer's confidence about the purchase decision (Schwenk, 1986; Tversky & Kahneman, 1974) and mitigate product-related uncertainty. Indeed, such predictions about the positive effects of review length have frequently been supported by evidence (Baek et al., 2012; Chevalier & Mayzlin, 2006; Hu & Chen, 2016; Korfiatis, García-Bariocanal, & Sánchez-Alonso, 2012; Pan & Zhang, 2011; Willemsen, Neijens, Bronner, & Ridder, 2011; Wu, 2013; Zhang, Craciun, & Shin, 2010).

Notwithstanding these positive effects, it follows from information processing theory that, at certain review lengths, information overload

* Corresponding author at: 1 Ben-Gurion Ave., Beer-Sheva 84105, Israel.
E-mail address: finkl@bgu.ac.il (L. Fink).

https://doi.org/10.1016/j.ijinfomgt.2017.11.002
Received 24 May 2017; Accepted 4 November 2017

may become an issue and the marginal utility of product-related information may become negative. Past the point of optimum cognitive load, information use may decline and render additional information ineffective and even detrimental to decision making. If, consistent with cognitive accounts of information processing, information contained in product reviews indeed has diminishing marginal utility, then the relationship between review length and effectiveness measures should be curvilinear and take the shape of an inverted U.

In this study, we argue that the positive effects attributed in the existing literature to review length have been the consequence of the frequent reliance on data that reflect the use of personal computers (PCs) rather than mobile devices to engage in electronic commerce. Because the PC setting is typically less constrained than the mobile setting in terms of the time, attention, and screen size available to users, larger amounts of information can be consumed by PC users before they experience information overload. In such settings, therefore, the point of optimum cognitive load, above which the marginal utility of product-related information is negative, is less likely to be observed, resulting in a conclusion that longer reviews are generally more beneficial to users. We hypothesize that in settings of mobile use, in which use is typically more constrained either by the environment or by features of the platform, the point of optimum cognitive load is more likely to be observed in data reflecting the behavior of users, resulting in a conclusion that longer reviews are not necessarily better.

We hypothesize that review length has a negative curvilinear relationship with product demand and test this hypothesis by collecting and analyzing large-scale data from two major stores for Android apps, Google Play and Amazon Appstore. We choose this specific research setting because consumers interact with mobile app stores primarily through mobile devices. This specific setting also allows the comparison of different ranges of review lengths and different product types with different pricing mechanisms. In particular, we collect four sets of data, about free and paid apps from Google Play and about free and paid apps from Amazon Appstore. For each app, we collect data about its demand (i.e., download measure) and about various app and review characteristics. We then regress app demand both on review length and on its squared term, while controlling for app and review characteristics. Across all four datasets, we find a statistically significant, negative coefficient for the squared review length, confirming that the relationship between review length and app demand indeed takes the shape of an inverted-U.

Our results contribute to the emerging literature on the effectiveness of online consumer-generated reviews by offering an explanation that accommodates conflicting findings about the effects of review length, while being reflective of theoretical accounts of information consumption that consider both information scarcity and overload. The findings have important implications for designers of electronic markets, who may wish to constrain the length of reviews in specific settings to increase their effectiveness.

2. Background and hypothesis

The literature generally considers review length to have a positive effect on various performance dimensions (Baek et al., 2012; Chevalier & Mayzlin, 2006; Hu & Chen, 2016; Korfiatis et al., 2012; Pan & Zhang, 2011; Willemsen et al., 2011; Wu, 2013; Zhang et al., 2010). A question that naturally follows is why previous studies described a positive relationship rather than a curvilinear one. The answer to this question is found in the characteristics of the research setting examined in previous work (e.g., Chevalier & Mayzlin, 2006; Floyd et al., 2014; Ghose & Ipeirotis, 2011) and in the lack of attention to idiosyncratic findings that are inconsistent with predictions (e.g., Fang et al., 2013; Huang, Chen, Yen, & Tran, 2015; Kuan et al., 2015).

As for previous work, until recent years, electronic commerce had typically been conducted by using a PC, which places few constraints in terms of time, attention, or screen size on the ability of consumers to process product-related information. In such a setting, large amounts of information can be processed before the consumer experiences information overload. In recent years, electronic commerce has rapidly migrated to mobile platforms, such as smartphones and tablet computers (Piccoli & Ott, 2014; Pousttchi, Tilson, Lyytinen, & Hufenbach, 2015; Wang, Malthouse, & Krishnamurthi, 2015). The mobile setting is more constrained than the PC setting in two aspects. First, by definition, mobile devices are more likely to be used when people are mobile (e.g., walking on the street or sitting in a train station) and can devote less attention to information processing than when they are using stationary PCs. Given that attention is the scarce resource in decision making (Simon, 1978), the need to pay more attention to the environment reduces the attention available for information processing. Second, screen sizes are typically smaller in the mobile setting than in the PC setting, leading to higher costs of information search (Ghose, Goldfarb, & Han, 2013) and to inferior task performance (Sweeney & Crestani, 2006). These two factors – lower attention to information processing and higher information search costs – which characterize the mobile setting in comparison to the PC setting, imply that information overload is likely to be experienced in the mobile setting after processing smaller amounts of information. Because previous research used data largely reflecting the PC setting, characterized by higher optimum load thresholds, it is reasonable that most review length observations were distributed at values lower than these thresholds, indicating that utility always increases with review length. This argument suggests that data reflecting the more constrained mobile setting, characterized by lower optimum load thresholds, is likely to include more review length observations that surpass these thresholds, providing evidence that the marginal utility of review length can be negative.

As for idiosyncratic findings, the existing literature, which has focused on information processing via PCs, does include a few accounts of contradicting evidence, which allude to the existence of a negative curvilinear effect of review length. Importantly, these studies hypothesize about a positive effect of review length, but their evidence suggests that this is not necessarily the case. Huang et al. (2015) examine the effect of review length on review helpfulness by analyzing data on six products from Amazon.com. They find that while review length has a significant positive effect for all reviews, this aggregate positive effect masks a positive effect for reviews shorter than average (144 words) and no effect for reviews longer than average. Based on these findings, they conclude that word count has a threshold in its effect on review helpfulness, above which “its effect diminishes significantly or becomes near non-existent” (Huang et al., 2015, p. 17). The same relationship is implicitly observed in an experiment by Lin, Huang, and Yang (2007), who hypothesize that review length positively affects consumer purchase intention, but find purchase intention to be lower for short reviews than for average and long reviews, without significant differences between consumers shown average and long reviews. Fang et al. (2013) offer several explanations for their finding of a negative effect of review length on product sales (despite their hypothesis of a positive effect), among which is the argument that longer reviews take more time to process. Finally, Kuan et al. (2015) hypothesize that various review attributes positively affect review helpfulness. They analyze reviews for two different products – DVDs and books – and find that review length has different nonlinear effects for different products.

While the idiosyncratic findings reviewed above are considered in the existing literature as inconsistent with predictions of positive effects of review length, these findings suggest that *descriptive* theory about information overload is as crucial to understanding the effects of review length as *normative* theory about information value. The present study thus aims at advancing research on online product reviews by providing a more nuanced account, theoretically anchored in models of human information processing, of the implications of review length. Consistent with the literature (Chevalier & Mayzlin, 2006; Floyd et al., 2014; Gu et al., 2012), we operationally define product effectiveness as product demand (i.e., actual sales). On the basis of the theoretical reasoning presented thus far in this section, we hypothesize that review length has a negative curvilinear relationship with product demand, which can be empirically observed when consumers process product-related information via mobile devices.

3. Method

We empirically tested our hypothesis about a negative curvilinear relationship of review length with product demand by collecting large-scale data from two major stores for Android apps, Google Play and Amazon Appstore. Mobile apps, the products offered in these stores, are acquired primarily through the more cognitively-constrained mobile setting, about which relatively little research has been conducted. For each store, we collected data about free and paid apps separately because of their distinguished characteristics (Liu, Au, & Choi, 2014; Viennot, Garcia, & Nieh, 2014) and because the stores independently rank free and paid apps. At the time of data collection (December 2013), the apps on Google Play were listed in 26 different categories, with the games category further divided into 18 sub-categories. For each category, separate lists of free and paid apps each included the top 540 apps (other apps were available only by using the search option) according to the number of downloads and to such other factors as rating, age, and history of growth (Liu et al., 2014). Amazon Appstore offered apps in 28 different categories, with the games category divided into 16 sub-categories. This store offered a more limited supply of apps and therefore its lists in the various categories included all available apps and not just the more popular ones.

We used software agents to automatically collect comprehensive data about all apps populating the ranking lists for the various categories on both stores. For each of these apps (the unit of analysis in this study), we collected data on the following app characteristics: price (in US Dollars, only for paid apps), age (number of days elapsed since the last update or since the app was released for Google or Amazon respectively), size (in MB), and a dummy variable indicating whether or not the app was a game. We also collected data for each app on the following review characteristics: average star rating (on the 1–5 star rating), number of reviews, and average review length (in number of characters).

Data collection yielded data on 33,119 apps on Google and 95,683 apps on Amazon. Consistent with the long-tail distribution characterizing demand in online markets (Brynjolfsson, Hu, & Simester, 2011; Garg & Telang, 2013), many of the apps had no reviews. The data included 23,850 apps on Google (12,065 free apps and 11,785 paid apps) and 27,248 apps on Amazon (16,757 free apps and 10,491 paid apps) with at least a single review. To avoid potential biases, consistent with approaches described in the literature (Ghose & Ipeirotis, 2011; Liu et al., 2014; Mudambi & Schuff, 2010), we removed data on apps for which there were few (three or less) reviews or the reviews included no text, yielding complete data on 7864 Google Free apps, 6206 Google Paid apps, 6158 Amazon Free apps, and 2734 Amazon Paid apps.

Different measures of product demand were used for the two stores. Because lists on Google are category-specific and take into account factors other than number of downloads, these lists could not be used for ranking purposes (Liu et al., 2014). The Google Play intervals for number of installs were consequently used to proxy for product demand. Our approach to data analysis was to separately analyze the datasets for free and paid apps, especially as Google uses two separate lists and Amazon ranks them independently. Data about demand for Google apps were therefore separated into two distinct distributions for number of downloads. Because demand for free apps is considerably higher than that for paid apps, their distributions were normalized differently (Table 1). The literature provides more guidance on how to measure product demand on Amazon. Studies commonly use the log of the Amazon sales rank as a proxy for the log of actual sales, based on findings that the distribution of actual sales in terms of sales rank has a Pareto distribution (Brynjolfsson, Hu, & Smith, 2003; Chevalier & Goolsbee, 2003; Garg & Telang, 2013). The negative log values of the Amazon sales ranks for free and paid apps were therefore used as measures of product demand. Negative values were used so that higher values on dependent variables reflected higher demand, consistent with the Google datasets. Importantly, data were collected from two stores to increase the validity of the findings by conducting two independent analyses. Because there was no need to integrate the data collected from the two stores, we could use different measures of product demand for different stores.

Table 1
Grouping of Google Play download intervals into download categories.

Download category	1	2	3	4	5	6	7
Google Free (N = 7864)							
1–5	0	0	0	0	0	0	0
5–10	0	0	0	0	0	0	0
10–50	1	0	0	0	0	0	0
50–100	1	0	0	0	0	0	0
100–500	10	0	0	0	0	0	0
500–1000	19	0	0	0	0	0	0
1000–5000	244	0	0	0	0	0	0
5000–10,000	0	292	0	0	0	0	0
10,000–50,000	0	1361	0	0	0	0	0
50,000–100,000	0	0	917	0	0	0	0
100,000–500,000	0	0	0	2391	0	0	0
500,000–1,000,000	0	0	0	0	889	0	0
1,000,000–5,000,000	0	0	0	0	0	1296	0
5,000,000–10,000,000	0	0	0	0	0	0	253
10,000,000–50,000,000	0	0	0	0	0	0	172
50,000,000–100,000,000	0	0	0	0	0	0	11
100,000,000–500,000,000	0	0	0	0	0	0	7
Total	275	1653	917	2391	889	1296	443
Google Paid (N = 6206)							
1–5	1	0	0	0	0	0	0
5–10	0	0	0	0	0	0	0
10–50	7	0	0	0	0	0	0
50–100	34	0	0	0	0	0	0
100–500	0	616	0	0	0	0	0
500–1000	0	0	797	0	0	0	0
1000–5000	0	0	0	2604	0	0	0
5000–10,000	0	0	0	0	838	0	0
10,000–50,000	0	0	0	0	0	1016	0
50,000–100,000	0	0	0	0	0	0	156
100,000–500,000	0	0	0	0	0	0	119
500,000–1,000,000	0	0	0	0	0	0	15
1,000,000–5,000,000	0	0	0	0	0	0	2
5,000,000–10,000,000	0	0	0	0	0	0	1
10,000,000–50,000,000	0	0	0	0	0	0	0
50,000,000–100,000,000	0	0	0	0	0	0	0
100,000,000–500,000,000	0	0	0	0	0	0	0
Total	42	616	797	2604	838	1016	293

App numbers are shown.

4. Results

Descriptive statistics for all variables included in data analysis and the correlation matrices are presented in Tables 2 and 3, respectively. Four linear specifications, presented in Eqs. (1) through (4), were used to analyze the Google Free (GF), Google Paid (GP), Amazon Free (AF), and Amazon Paid (AP) datasets.

$$\begin{aligned}\text{Downloads}_i = {} & \beta_{0,GF} + \beta_{2,GF}\cdot\log(\text{Age}_i) + \beta_{3,GF}\cdot\log(\text{Size}_i) + \beta_{4,GF} \\ & \cdot\text{Game}_i + \beta_{5,GF}\cdot\text{StarRating}_i + \beta_{6,GF}\cdot\text{ReviewNumber}_i + \beta_{7,GF} \\ & \cdot\log(\text{ReviewLength}_i) + \beta_{8,GF}\cdot(\log(\text{ReviewLength}_i))^2 + \varepsilon_{GF}\end{aligned} \tag{1}$$

Table 2
Descriptive statistics for mobile apps.

	Mean	Standard deviation	Min	Max
Google Free (N = 7864)				
Download category	3.97	1.60	1	7
Days since update	132.85	196.25	3	1617
App size	9.98	30.78	0.01	1300
Game	0.28	0.45	0	1
Star rating	4.14	0.45	1.50	5.00
Number of reviews	10672.50	82446.84	5	5088951
Average review length	85.39	54.05	10.25	699.00
Google Paid (N = 6206)				
Download category	4.26	1.32	1	7
Price	3.69	7.04	0.85	199.71
Days since update	218.40	284.33	4	1701
App size	16.65	56.29	0.01	1300
Game	0.22	0.41	0	1
Star rating	4.12	0.60	1.00	5.00
Number of reviews	430.91	2393.43	4	108728
Average review length	148.99	66.80	13.00	831.33
Amazon Free (N = 6158)				
Sales rank	11777.39	10265.59	2	70304
Days since release	524.11	278.60	9	1079
App size	15.16	41.77	0.01	1500
Game	0.39	0.49	0	1
Star rating	3.40	0.87	1.00	5.00
Number of reviews	115.79	685.29	4	24978
Average review length	196.49	77.13	2.00	1454.42
Amazon Paid (N = 2734)				
Sales rank	6565.15	5727.54	1	39583
Price	2.39	4.57	0.99	99.00
Days since release	566.04	267.68	1	1019
App size	26.51	70.75	0.02	1400
Game	0.49	0.50	0	1
Star rating	3.43	0.92	1.00	5.00
Number of reviews	73.02	342.25	4	14251
Average review length	239.11	108.48	8.00	1248.00

Price is in US Dollars (only for paid apps). App size is in MB. Game is a dummy variable indicating whether the app is a game (1) or not (0). Star rating is the average star rating on the 1–5 scale. Average review length is in number of characters.

$$\begin{aligned}\text{Downloads}_i = {} & \beta_{0,GP} + \beta_{1,\text{GP}}\cdot\log(\text{Price}_i) + \beta_{2,\text{GP}}\cdot\log(\text{Age}_i) + \beta_{3,\text{GP}} \\ & \cdot\log(\text{Size}_i) + \beta_{4,\text{GP}}\cdot\text{Game}_i + \beta_{5,\text{GP}}\cdot\text{StarRating}_i + \beta_{6,\text{GP}} \\ & \cdot\text{ReviewNumber}_i + \beta_{7,\text{GP}}\cdot\log(\text{ReviewLength}_i) + \beta_{8,\text{GP}} \\ & \cdot(\log(\text{ReviewLength}_i))^2 + \epsilon_{\text{GP}}\end{aligned} \quad (2)$$

$$\begin{aligned}-\log(\text{SalesRank}_i) = {} & \beta_{0,AF} + \beta_{2,\text{AF}}\cdot\log(\text{Age}_i) + \beta_{3,\text{AF}}\cdot\log(\text{Size}_i) + \beta_{4,\text{AF}} \\ & \cdot\text{Game}_i + \beta_{5,\text{AF}}\cdot\text{StarRating}_i + \beta_{6,\text{AF}} \\ & \cdot\text{ReviewNumber}_i + \beta_{7,\text{AF}}\cdot\log(\text{ReviewLength}_i) + \beta_{8,\text{AF}} \\ & \cdot(\log(\text{ReviewLength}_i))^2 + \epsilon_{\text{AF}}\end{aligned} \quad (3)$$

$$\begin{aligned}-\log(\text{SalesRank}_i) = {} & \beta_{0,\text{AP}} + \beta_{1,\text{AP}}\cdot\log(\text{Price}_i) + \beta_{2,\text{AP}} \\ & \cdot\log(\text{Age}_i) + \beta_{3,\text{AP}}\cdot\log(\text{Size}_i) + \beta_{4,\text{AP}}\cdot\text{Game}_i + \beta_{5,\text{AP}} \\ & \cdot\text{StarRating}_i + \beta_{6,\text{AP}}\cdot\text{ReviewNumber}_i + \beta_{7,\text{AP}} \\ & \cdot\log(\text{ReviewLength}_i) + \beta_{8,\text{AP}}\cdot(\log(\text{ReviewLength}_i))^2 \\ & + \epsilon_{\text{AP}}\end{aligned} \quad (4)$$

In the two specifications for Google, Eqs. (1) and (2), the dependent variable $Downloads_i$ was the respective download category for app i, whereas in the two specifications for Amazon, Eqs. (3) and (4), the term $-\log(SalesRank_i)$ proxied for the number of downloads of app i. As noted earlier, higher values on both dependent variables represented larger numbers of downloads. Within each store, the difference between the specifications for free and paid apps was the inclusion of a term for app price.

The specifications included linear and squared terms of the log of review length, while controlling for the various app and review characteristics. The log of review length was centered with respect to its mean to reduce the threat of multicollinearity that might arise when estimating models that include multiple terms with the same variable (Aiken & West, 1991; Jaccard & Turrisi, 2003). Multicollinearity was not a problem in any analysis, as evident by all variance inflation factors being below 1.5. The resulting estimates are presented in Table 4 and visually described in Fig. 1. The coefficients for size, star rating, and number of reviews were significantly positive for all four datasets. The coefficient for the game dummy was significantly positive for all datasets except for Amazon Paid, indicating that this was the only dataset in which game apps were not in higher demand compared to other apps. While the coefficients for age were nonsignificant for both Google datasets, these coefficients were significantly negative for both Amazon datasets, probably the consequence of age being defined differently across stores (number of days elapsed since update on Google and since release on Amazon). Similarly, while the coefficient for price was nonsignificant for Google, it was significantly negative for Amazon, suggesting that more expensive apps are in lower demand on Amazon but not on Google.

Most importantly, the coefficient for squared review length was significantly negative across all four datasets, providing consistent evidence of a negative curvilinear relationship between review length and product demand. This finding was insensitive to model re-specifications, including the removal of control variables, in all datasets.

Also of interest were the coefficients for the linear term of review length. This coefficient was significantly positive for Google Free, nonsignificant for Amazon Free, and significantly negative for both Google Paid and Amazon Paid. This finding suggested that the four datasets were distributed differently on the theoretical inverted-U-shaped curve. We tested this supposition by estimating the value of review length for which demand was highest, corresponding to the threshold of optimum cognitive load, after which the marginal utility of review length became negative. This was accomplished by calculating the review length for which the first derivative of quadratic Eqs. (1) through (4) equaled zero. The first derivative of Eqs. (1) and (2) for Google is presented in Eq. (5), and the first derivative of Eqs. (3) and (4) for Amazon is presented in Eq. (6).

$$\frac{\partial Downloads}{\partial \log(ReviewLength)} = \beta_{7,GF/GP} + 2\cdot\beta_{8,GF/GP}\cdot\log(ReviewLength) = 0 \quad (5)$$

$$\frac{\partial(-\log(SalesRank))}{\partial \log(ReviewLength)} = \beta_{7,AF/AP} + 2\cdot\beta_{8,AF/AP}\cdot\log(ReviewLength) = 0 \quad (6)$$

The resulting values were then rescaled (to reverse the mean-centralization) and exponetiated to yield the review lengths for which demand was highest. These values (in number of characters) were 128.85 for Google Free, 173.98 for Amazon Free, 114.88 for Google Paid, and 112.45 for Amazon Paid, indicating that maximum demand was reached at lower review lengths for paid apps compared to free apps. We could furthermore compare the mean and maximum-demand review lengths for each dataset to ascertain whether marginal utility of review length for the majority of observations was positive (ascending part of the curve), zero (top of the curve), or negative (descending part of the curve). Fig. 2 depicts predicted demand as a function of review length (based on the estimation of our specifications) against the observed distribution of review length for all four datasets. This analysis confirmed that mean review length was lower than maximum-demand review length for Google Free, relatively similar to it for Amazon Free, and higher than it for Google Paid and Amazon Paid.

Table 3
Correlation matrices.

	Demand	Game	Age	Size	Star rating	Review number	Review length	Review length^2
Google Free								
Demand	1							
Game	0.21***	1						
Age	−0.05***	−0.17***	1					
Size	0.23***	0.43***	−0.31***	1				
Star rating	0.08***	−0.03*	−0.09***	0.10***	1			
Rev. number	0.20***	0.04***	−0.06***	0.07***	0.07***	1		
Rev. length	0.14***	−0.19***	0.07***	−0.04**	−0.01	0.12***	1	
Rev. length^2	−0.08***	0.05***	−0.05***	−0.01	−0.04**	0.07***	−0.17***	1
Google Paid								
Demand	1							
Game	0.16***	1						
Age	−0.03*	−0.01	1					
Size	0.09***	0.36***	−0.24***	1				
Star rating	0.08***	−0.02	−0.21***	0.01	1			
Rev. number	0.28***	0.06***	−0.07***	0.02	0.07***	1		
Rev. length	−0.08***	−0.13***	−0.11***	0	−0.08***	0.04**	1	
Rev. length^2	−0.18***	−0.01	−0.03*	0.02	−0.02	−0.02	−0.07***	1
Price	−0.04***	−0.09***	−0.09***	0.07***	−0.05***	0	0.17***	0.08***
Amazon Free								
Demand	1							
Game	0.28***	1						
Age	−0.42***	−0.23***	1					
Size	0.39***	0.46***	−0.36***	1				
Star rating	0.23***	0.04**	−0.10***	0.18***	1			
Rev. number	0.31***	0.12***	0.02	0.14***	0.14***	1		
Rev. length	−0.03*	−0.10***	0.06***	0.01	0.08***	0.01	1	
Rev. length^2	−0.07***	−0.02	0	0	−0.02	−0.02	0.17***	1
Amazon Paid								
Demand	1							
Game	0.12***	1						
Age	−0.37***	−0.20***	1					
Size	0.25***	0.46***	−0.30***	1				
Star rating	0.22***	−0.05**	−0.03	0.15***	1			
Rev. number	0.32***	0.09***	0.05*	0.08***	0.10***	1		
Rev. length	−0.07***	−0.11***	0.02	0	0.17***	0.02	1	
Rev. length^2	−0.09***	−0.01	0	−0.02	−0.02	−0.04*	−0.09***	1
Price	−0.08***	−0.09***	0.03	0	0.02	0.01	0.13***	0.06**

Pearson correlation coefficients are shown. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; two-tailed p values are reported. App demand is the download category for Google and minus natural log of sales rank for Amazon. Price is in US Dollars (only for paid apps). Age is the natural log of the number of days elapsed since the last update or since the app was released for Google or Amazon respectively. Size is the natural log of app size in MB. Game is a dummy variable indicating whether the app is a game (1) or not (0). Star rating is the average star rating on the 1-5 scale. Review number is the number of reviews in thousands. Review length is the mean-centered natural log of the average review length in number of characters.

5. Discussion

The finding that maximum demand, which we interpret as corresponding to optimum cognitive load, is reached at lower values of review length for paid apps than for free apps proposes that the utilization of cognitive resources is contingent on product type, in our case on whether or not the product is acquired at a zero price. Characteristic of both stores, information overload seems to be associated with less information when actual payment is required. A reasonable explanation for this finding draws on frameworks that view motivation as consequential for information processing. One such framework is the elaboration likelihood model (Petty & Cacioppo, 1986), which describes two routes to persuasion, a central route and a peripheral route, where the specific route through which the persuasive message is processed depends on the motivation and ability of the recipient to process the message. When the recipient is motivated and able to process the message, considerable cognitive resources are expended to systematically process it through the central route. When the recipient is unmotivated or unable to process the message, little cognitive resources are expended to heuristically process it via the peripheral route. Consumers of paid apps should be more motivated to process app-related information than consumers of free apps because of the irreversibility of the purchase of a paid app against the ability to easily uninstall a free app with no financial ramifications. Paid app consumers are therefore likely to invest more cognitive resources in information processing and read the reviews more attentively compared to free app consumers. Consequently, free app consumers may consume more information than paid app consumers before experiencing information overload.

These different modes of information processing for free and paid apps have implications not only for theoretical cognitive thresholds but also for actual behavior. We observe lower review lengths for free apps than for paid apps. This observation is consistent with information processing motivation being lower for free app consumers, who may require less information to reach the decision threshold. Understanding of consumer behavior can therefore be enhanced by comparing observed and optimal review lengths. As can be seen in Fig. 2, review length observations are mostly found on the ascending (Google) or flat (Amazon) parts of the fitted curves for free apps and on the descending parts of the fitted curves for paid apps. Put differently, the decision threshold is likely to precede the information overload threshold for free apps and exceed it for paid apps, implying that consumers of paid apps, as a result of attributing higher importance to information processing, are more likely than consumers of free apps to experience information overload in the decision making process.

The finding that review length has a negative curvilinear relationship with demand for both free and paid apps on two major stores strengthens the external validity of our results and the confidence that they are not an artifact of overfitting. The results serve as another

Table 4
Results of OLS regressions of app demand on app and review characteristics.

Variables	Google Play		Amazon Appstore	
	Free	Paid	Free	Paid
Intercept	2.535*** (0.170)	3.766*** (0.136)	−6.224*** (0.138)	−5.341*** (0.214)
Price		−0.0007 (0.002)		−0.018*** (0.004)
Age	0.014 (0.013)	−0.004 (0.012)	−0.589*** (0.020)	−0.627*** (0.031)
Size	0.159*** (0.013)	0.027** (0.009)	0.136*** (0.010)	0.079*** (0.013)
Game	0.702*** (0.042)	0.383*** (0.041)	0.213*** (0.031)	−0.085 (0.047)
Star rating	0.253*** (0.037)	0.115*** (0.027)	0.192*** (0.016)	0.239*** (0.023)
Review number	0.003*** (0.0002)	0.150*** (0.007)	0.507*** (0.020)	1.143*** (0.060)
Review length	0.333*** (0.029)	−0.275*** (0.038)	−0.032 (0.044)	−0.307*** (0.052)
Review length^2	−0.284*** (0.035)	−0.801*** (0.051)	−0.238*** (0.042)	−0.233*** (0.051)
R^2	0.148	0.146	0.351	0.304
Adjusted R^2	0.147	0.145	0.350	0.302
F	195.4***	132.9***	474.6***	149.0***
df	7, 7856	8,6197	7,6150	8,2725

Unstandardized coefficients are shown, with standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; two-tailed p values are reported. App demand is the download category for Google and minus natural log of sales rank for Amazon. Price is in US Dollars (only for paid apps). Age is the natural log of the number of days elapsed since the last update or since the app was released for Google or Amazon respectively. Size is the natural log of app size in MB. Game is a dummy variable indicating whether the app is a game (1) or not (0). Star rating is the average star rating on the 1–5 scale. Review number is the number of reviews in thousands. Review length is the mean-centered natural log of the average review length in number of characters.

demonstration of the non-linear and non-monotonous relationship between the availability of information and decision making in an important and innovative setting that the research community has only recently begun exploring. This study advances research on online consumer behavior by showing that perceptions of longer online reviews as necessarily better should be reconsidered given our findings of nonlinear relationships between review length and product demand in the mobile setting. In so doing, the study provides an explanation that reconciles existing normative predictions with descriptive decision theory. Designers of electronic platforms and markets should be aware that cognitive processes underlying consumer behavior are not necessarily reflective of monotonous relationships between information availability and effectiveness. Restrictions on information availability in electronic commerce may therefore be warranted under certain circumstances. Future research should follow the approach taken in the present study and focus on specific characteristics of online product reviews to examine whether their effects are more intricate than previously postulated. Information processing theory and related frameworks, such as the elaboration likelihood model, provide the theoretical infrastructure upon which such effects of specific characteristics can be empirically investigated.

6. Conclusion

In this study, we focus on review length among the numerous characteristics of online consumer-generated reviews and investigate whether its effects conform to information processing theory and to predictions about the performance-related consequences of information overload. In particular, we test whether review length has a negative curvilinear relationship with product demand on app stores. We argue that the reason such a curvilinear relationship has not been frequently observed in previous studies is their analysis of data about consumer behavior on PCs, which afford a less constrained environment in terms of information processing relative to mobile devices. Because the environment is less constrained, the point of optimum cognitive load, above which the marginal value of review length is negative, is less likely to be observed empirically. With the ubiquity of mobile devices and the popularity of mobile apps comes an opportunity to empirically observe this point of optimum load in electronic commerce. We address this opportunity by collecting large-scale data on free and paid apps on both Google Play and Amazon Appstore. Four specifications are then formulated to estimate the effects of both linear and squared review

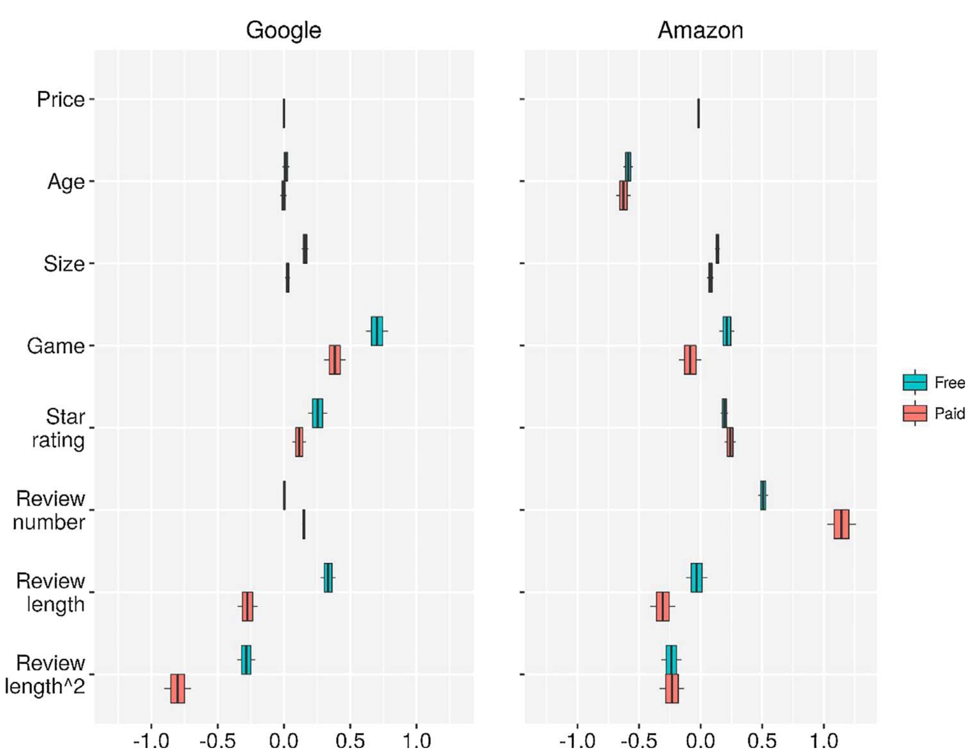

Fig. 1. Visualization of OLS regression results (Table 4). Unstandardized coefficients are presented for Google Play (A) and Amazon Appstore (B), with standard errors (boxes) and 95% confidence intervals (whiskers).

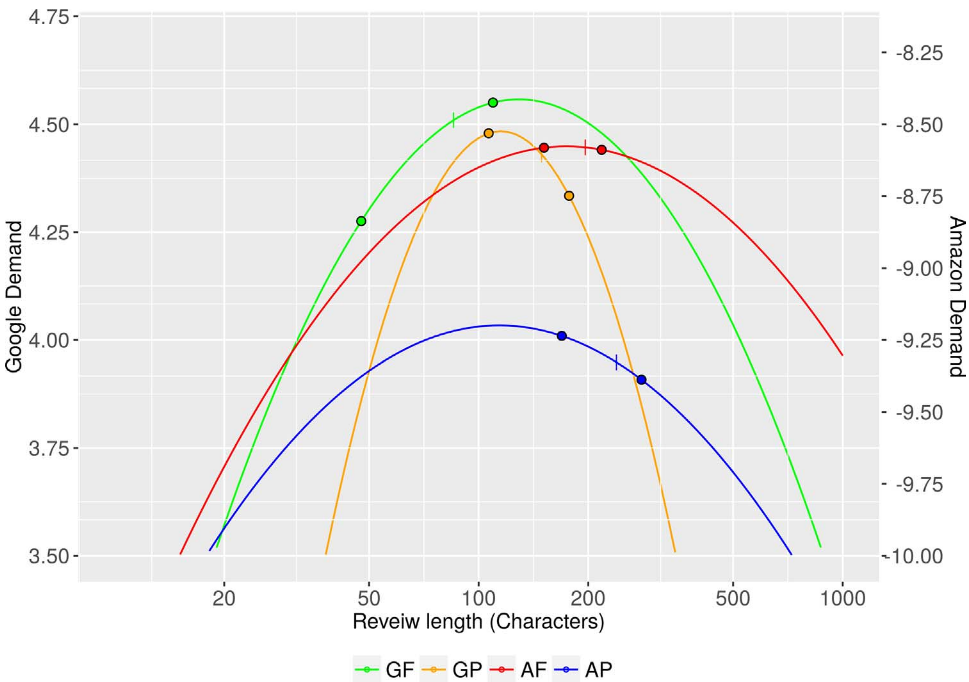

Fig. 2. Fitted curves. Fitted curves of predicted demand as a function of review length based on the estimation of our specifications with the data for Google Free (GF), Google Paid (GP), Amazon Free (AF) and Amazon Paid (AP). The observed distributions of review length are presented as means (ticks) and one quartile ranges (circles). The scale for demand on Google (left y-axis) is the Google download category, and the scale for demand on Amazon (right y-axis) is minus natural log of sales rank.

length on product demand, while controlling for app and review characteristics. We find the coefficients of squared review length to be negative for all four datasets, implying that review length has a negative curvilinear relationship with demand for free and paid apps on both app stores. We furthermore find that the optimum load points, where the relationship changes from positive to negative (zero marginal utility of review length), are higher for free apps than for paid apps. Another interesting finding is that the observed distributions of review length for free apps mostly lie on the ascending or flat parts of the curve, whereas these distributions for paid apps lie on the descending parts of the curve. We interpret these findings as supporting the theoretical notion that consumers of free apps are less motivated than consumers of paid apps to process product-related information. Overall, our findings suggest that settings of mobile use may be more suitable than settings of PC use to study the implications of information overload on the effectiveness of online product reviews.

References

Ahmad, S. N., & Laroche, M. (2017). Analyzing electronic word of mouth: A social commerce construct. *International Journal of Information Management, 37*(3), 202–213.
Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Thousand Oaks, CA: Sage.
Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce, 17*(2), 99–126.
Brynjolfsson, E., Hu, Y., & Smith, M. D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science, 49*(11), 1580–1596.
Brynjolfsson, E., Hu, Y., & Simester, D. (2011). Goodbye pareto principle, hello long tail: The effect of search costs on the concentration of product sales. *Management Science, 57*(8), 1373–1386.
Chevalier, J., & Goolsbee, A. (2003). Measuring prices and price competition online: Amazon.com and Barnesandnoble.com. *Quantitative Marketing and Economics, 1*(2), 203–222.
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.
Driver, M. J., & Mock, T. J. (1975). Human information processing, decision style theory, and accounting information systems. *The Accounting Review, 50*(3), 490–508.
Fang, H., Zhang, J., Bao, Y., & Zhu, Q. (2013). Towards effective online review systems in the Chinese context: A cross-cultural empirical study. *Electronic Commerce Research and Applications, 12*(3), 208–220.
Floyd, K., Freling, R., Alhoqail, S., Cho, H. Y., & Freling, T. (2014). How online product reviews affect retail sales: A meta-analysis. *Journal of Retailing, 90*(2), 217–232.
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.
Garg, R., & Telang, R. (2013). Inferring app demand from publicly available data. *MIS Quarterly, 37*(4), 1253–1264.
Ghose, A., & Ipeirotis, P. G. (2011). 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.
Ghose, A., Goldfarb, A., & Han, S. P. (2013). How is the mobile internet different? Search costs and local activities. *Information Systems Research, 24*(3), 613–631.
Gross, B. M. (1964). *The managing of organizations: The administrative struggle.* New York: Free Press of Glencoe.
Gu, B., Park, J., & Konana, P. (2012). The impact of external word-of-mouth sources on retailer sales of high-involvement products. *Information Systems Research, 23*(1), 182–196.
Hu, Y.-H., & Chen, K. (2016). Predicting hotel review helpfulness: The impact of review visibility, and interaction between hotel stars and review ratings. *International Journal of Information Management, 36*(6), 929–944.
Huang, A. H., Chen, K., Yen, D. C., & Tran, T. P. (2015). A study of factors that contribute to online review helpfulness. *Computers in Human Behavior, 48*, 17–27.
Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression* (2nd ed.). Thousand Oaks, CA: Sage.
Jacoby, J., Speller, D. E., & Kohn, C. A. (1974). Brand choice behavior as a function of information load. *Journal of Marketing Research, 11*(1), 63–69.
Jacoby, J. (1977). Information load and decision quality: Some contested issues. *Journal of Marketing Research, 14*(4), 569–573.
Jones, Q., Ravid, G., & Rafaeli, S. (2004). Information overload and the message dynamics of online interaction spaces: A theoretical model and empirical exploration. *Information Systems Research, 15*(2), 194–210.
King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of Interactive Marketing, 28*(3), 167–183.
Korfiatis, N., García-Bariocanal, E., & Sánchez-Alonso, S. (2012). Evaluating content quality and helpfulness of online product reviews: The interplay of review helpfulness vs. review content. *Electronic Commerce Research and Applications, 11*(3), 205–217.
Kuan, K. K. Y., Hui, K. L., Prasarnphanich, P., & Lai, H. Y. (2015). What makes a review voted? An empirical investigation of review voting in online review systems. *Journal of the Association for Information Systems, 16*(1), 48–71.
Lin, T. M. Y., Huang, Y. K., & Yang, W. I. (2007). An experimental design approach to investigating the relationship between internet book reviews and purchase intention. *Library & Information Science Research, 29*(3), 397–415.
Liu, C. Z., Au, Y. A., & Choi, H. S. (2014). Effects of freemium strategy in the mobile app market: An empirical study of Google Play. *Journal of Management Information Systems, 31*(3), 326–354.
Malhotra, N. K. (1982). Information load and consumer decision making. *Journal of Consumer Research, 8*(4), 419–430.
Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly, 34*(1), 185–200.
Neirotti, P., Raguseo, E., & Paolucci, E. (2016). Are customers' reviews creating value in the hospitality industry? Exploring the moderating effects of market positioning. *International Journal of Information Management, 36*(6), 1133–1143.
O'Reilly, C. A. (1980). Individuals and information overload in organizations: Is more necessarily better? *Academy of Management Journal, 23*(4), 684–696.

Pan, Y., & Zhang, J. Q. (2011). Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing, 87*(4), 598–612.
Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Vol. Ed.), *Advances in experimental social psychology: Vol. 19*, (pp. 123–205). Orlando, FL: Academic Press.
Piccoli, G., & Ott, M. (2014). Impact of mobility and timing on user-generated content. *MIS Quarterly Executive, 13*(3), 147–157.
Pousttchi, K., Tilson, D., Lyytinen, K., & Hufenbach, Y. (2015). Mobile commerce research yesterday, today, tomorrow—What remains to be done? *International Journal of Electronic Commerce, 19*(4), 1–20.
Schroder, H., Driver, M., & Streufert, S. (1967). *Human information processing: Individual's and group's functioning in complex social situations.* New York: Holt, Rinehart and Winston.
Schwenk, C. R. (1986). Information, cognitive biases, and commitment to a course of action. *Academy of Management Review, 11*(2), 298–310.
Simon, H. A. (1978). Rationality as process and as product of thought. *American Economic Review, 68*(2), 1–16.
Sweeney, S., & Crestani, F. (2006). Effective search results summary size and device screen size: Is there a relationship? *Information Processing and Management, 42*(4), 1056–1074.
Toffler, A. (1970). *Future shock.* New York: Bantam Books.
Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131.
Viennot, N., Garcia, E., & Nieh, J. (2014). A measurement study of Google Play. *ACM SIGMETRICS Performance Evaluation Review, 42*(1), 221–233.
Wang, R. J. H., Malthouse, E. C., & Krishnamurthi, L. (2015). On the go: How mobile shopping affects customer purchase behavior. *Journal of Retailing, 91*(2), 217–234.
Willemsen, L. M., Neijens, P. C., Bronner, F., & Ridder, J. A. D. (2011). Highly recommended! The content characteristics and perceived usefulness of online consumer reviews. *Journal of Computer-Mediated Communication, 17*(1), 19–38.
Wu, P. F. (2013). In search of negativity bias: An empirical study of perceived helpfulness of online reviews. *Psychology and Marketing, 30*(11), 971–984.
Zhang, J. Q., Craciun, G., & Shin, D. (2010). When does electronic word-of-mouth matter? A study of consumer product reviews. *Journal of Business Research, 63*(12), 1336–1341.