AMERICAN PSYCHOLOGICAL ASSOCIATION

Decision

ISSN: 2325-9965

2024, Vol. 11, No. 4, 439–449
https://doi.org/10.1037/dec0000238

Attention-Driven Imitation in Consumer Reviews

Charles Alba[1, 2], Lukasz Walasek[1], and Mikhail S. Spektor[1]

[1] Department of Psychology, University of Warwick

[2] Division of Computational and Data Sciences, Washington University in St. Louis

Product reviews on e-commerce platforms can have a pronounced effect on consumers' decisions. Less is known, however, whether the reviews written by others can shape a person's own written opinion of a product. We hypothesized that people who compose reviews on digital storefronts will try to imitate successful reviews, such that their content will show similarity with other reviews displayed at the time of writing. More specifically, we predicted that reviews would be more semantically similar to the most successful, salient, and readily accessible reviews written by others. To investigate this hypothesis, we extracted over 3 million reviews from a major online distribution platform and traced the reviews that were displayed at the time when each review was being composed. Using word embeddings from a pretrained language model, we quantified the semantic similarity between a given review and other reviews that were visible (or not) to a user. We found that reviewers imitate the most helpful reviews written by others, especially those that are visually salient. Their reviews, in turn, gather more helpfulness ratings in the future, leading to a cascade of similar reviews. Our findings suggest that the default sorting and display format of reviews on online platforms will have a pronounced effect on the style and content of new reviews.

Keywords: online reviews, similarity, salience, order effects, text mining

Over the past few decades, the growing accessibility and popularity of digital storefronts have significantly changed how consumers make their purchasing decisions. A core feature of many modern online marketplaces is that they enable their customers to express opinions about their purchases by writing consumer reviews. It is now well-established that such electronic word-of-mouth communication can have a considerable impact on consumers' decisions (Chevalier & Mayzlin, 2006). Much of the existing research on the topic of electronic word-of-mouth focuses primarily on how specific features of written reviews (e.g., their style or content) and the unique characteristics of the online platforms (e.g., product types offered, design of the user interface) influence consumers' decision-making processes and purchasing behavior (e.g., as measured by sales, consumers' information search, or their purchase intention; see Babić Rosario et al., 2016; Liu et al., 2019, for recent comprehensive reviews). A comparatively smaller stream of research focuses on the reviewers themselves, investigating different motivations that drive people to contribute their opinions on e-commerce platforms. Research from this literature has shown that people may write reviews to punish a seller (Lafky, 2014) or simply to relive

This article was published Online First July 18, 2024.

Charles Alba https://orcid.org/0000-0001-7711-360X
Lukasz Walasek https://orcid.org/0000-0002-7360-0037
Mikhail S. Spektor https://orcid.org/0000-0003-0652-1993

Charles Alba played a lead role in data curation, formal analysis, investigation, methodology, and project administration, an equal role in conceptualization and writing–original draft, and a supporting role in visualization. Lukasz Walasek played an equal role in conceptualization, supervision, and writing–review and editing and a supporting role in formal analysis, methodology, and project administration. Mikhail S. Spektor played a lead role in visualization, a supporting role in formal analysis, methodology, and project administration, and an equal role in conceptualization, supervision, writing–original draft, and writing–review and editing.

Correspondence concerning this article should be addressed to Lukasz Walasek or Mikhail S. Spektor, Department of Psychology, University of Warwick, Coventry, CV4 7AL, United Kingdom. Email: l.walasek@warwick.ac.uk or mikhail@spektor.ch

their consumption experience (Yoo & Gretzel, 2008). At the same time, reviewers may also contribute in order to help other buyers make better decisions (Yoo & Gretzel, 2008) or to enhance their own feelings of belonging by being active members of a community (Cheung & Lee, 2012). These motivations are likely to influence reviewers' own decisions on the content and style of their own review (Schindler & Bickart, 2012). To be able to help other consumers make better decisions, reviewers need to decide what makes a useful review. How do they do this? Are the reviewers influenced by the opinions shared by other users? If so, do more salient reviews have more impact on people's own contributions? In the present article, we investigate whether reviewers' own contributions might be shaped by the immediate context afforded by the design features of the online platform. The goal of the present work is to understand the role this review-writing context plays in influencing the reviews that are written.

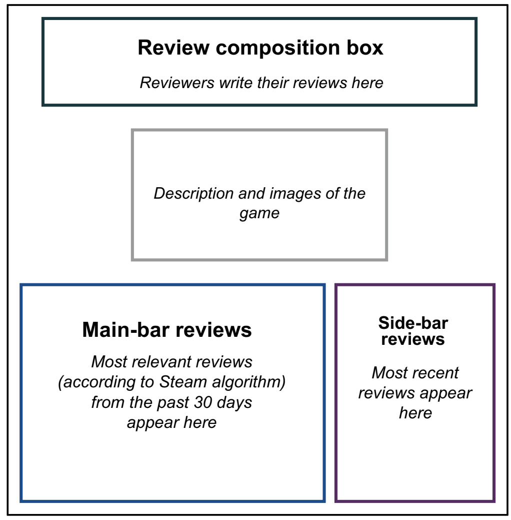
For this purpose, we build on a simple conceptual framework for understanding the diverse sources of influence on how consumers compose their reviews (see, e.g., Berger et al., 2020). Our fundamental assumption is that reviewers follow two broad types of motivation: First, every review is an expression of a person's own experience, and so their reviews are evaluative statements about that individual's consumption experience. Second, reviewers are also motivated by the goal of making a valued contribution to the community of other consumers on the same platform. Accordingly, they will craft their reviews with the goal of maximizing their helpfulness to others. Irrespective of whether the motivation to contribute on electronic word-of-mouth platforms fulfills personal needs (e.g., a sense of belonging among people with similar preferences) or reflects people's concerns with others' well-being (e.g., a sense of obligation to help others avoid bad products), it is likely that consumers contribute reviews by considering what types of reviews are valued in a given setting (Cheung & Lee, 2012).

Abstracting from a person's idiosyncratic consumption experience, what features of a review make it successful or liked by the community? In other words, how do consumers determine what makes a review "good"? Here, we propose that the main cognitive mechanism behind composing a good review is imitation (Offerman & Sonnemans, 1998). Consumers learn from the broader context of other available reviews about the desirable features of a good review. This mechanism could be deliberate, such that consumers actively try to mirror reviews that are rated as helpful (Eberhard et al., 2018), effectively constructing hypotheses about potential features of a review that could lead to a positive evaluation by others. If reviewers' perceptions were accurate (or they had in-depth knowledge of the academic literature), they could determine that reviews rated as helpful are those that include emotional language (Ahmad & Laroche, 2015), are more fluently written (Fang et al., 2016; Kronrod & Danziger, 2013; Moore, 2015; van Laer et al., 2018), are written by people who share broader consensus about a product (Naylor et al., 2011), or appear to have been written with more effort (Grewal & Stephen, 2019). Equally, consumers could avoid features of reviews that are too polarizing (Schoenmueller et al., 2020) or give the impression that they have been written by someone who did not purchase a product (Anderson & Simester, 2014). On a more implicit level, however, reviewers could also be influenced by the less relevant factors (e.g., Brandes & Dover, 2022). Here, a range of bottom-up attentional mechanisms could determine which reviews are imitated (Ashby et al., 2015). First, we expect that the visual salience should play a significant role. There is a large amount of literature in basic research showing that particularly salient elements capture decision-makers' attention (Desimone & Duncan, 1995), and there is evidence that larger and more central displays are more salient (Buscher et al., 2009; Kosslyn & Alper, 1977; Roth et al., 2013; Yantis & Jonides, 1984). Based on this research, we expect that reviewers pay more attention to, and therefore imitate, the more prominently featured reviews written by others. Second, but relatedly, the order in which reviews are displayed may influence which reviews are imitated as well. In terms of purchasing decisions, past research has shown that reviews that are displayed first (i.e., on top) have a larger effect than those shown thereafter (Kapoor & Piramuthu, 2009; Vana & Lambrecht, 2021; Wang et al., 2015). Evidence from eye-tracking studies also shows that reviews are scanned sequentially and that reviews that are later in a sequence are processed more superficially (McCarthy, 2013; Nielsen, 2010). In line with the rationale behind the effect of salience, we expect reviewers to be most affected by those that are displayed at the top of a page where they are readily accessible.

To investigate our hypothesis and assess to what extent reviewers imitate other people's reviews, we rely on the data from a large online video game platform. We developed an algorithm to trace back and simulate all the reviews that were visible to each reviewer at the time at which the reviewers were composing their own reviews. We use natural language processing to represent user-generated reviews in a vector space and compute the similarity between them. Foreshadowing our main results, we find that user reviews are most similar to prominently displayed and salient reviews, appearing at the top of the center of the page, and the more similar a review is to the most salient review displayed on the page, the more helpfulness votes it receives. Reviewers seem to attempt to imitate successful reviews, which in turn increases the chances of their reviews becoming successful themselves. Taken together, our results highlight the importance of display characteristics on user-generated content.

Figure 1

A Sketch of the Webpage That a Reviewer Would Observe While Writing Their Own Review on Steam



Note. Figure is for illustration purposes and is not drawn to scale. See the online article for the color version of this figure.

Method

The present study uses the "Steam" online video game distributional platform to investigate the cognitive processes underlying review writing. To do so, we computed the similarity between a large number of reviews on this platform and reviews that were accessible to the reviewers at the time of writing their reviews. The main advantage of Steam is that the user-generated reviews are presented in a rather unique manner that allows one to disentangle the cognitive processes discussed above (see Figure 1). Here, two types of reviews are shown in the review section of each video game page. The first is what we will refer to as the "main-bar reviews." These reviews make up a visually dominant proportion of the review section on a Steam webpage. They are sorted by helpfulness and explicitly presented to users as "MOST HELPFUL REVIEWS" (in capital letters). We classify reviews appearing there as being "salient." The second type of review consists of recency-sorted reviews. These reviews are displayed in a smaller sidebar of the review section and positioned next to the main bar on the right-most part of the page. Reviews appearing there are classified as being "non-salient." Additionally, we identified and extracted reviews that had been written within the previous 30 days (from the time when the target review was written) but which were *not* visible on either the main or the sidebar of the review page. These reviews are used to obtain a control condition. As the second independent variable, we coded the appearance of reviews in order (from top to bottom).

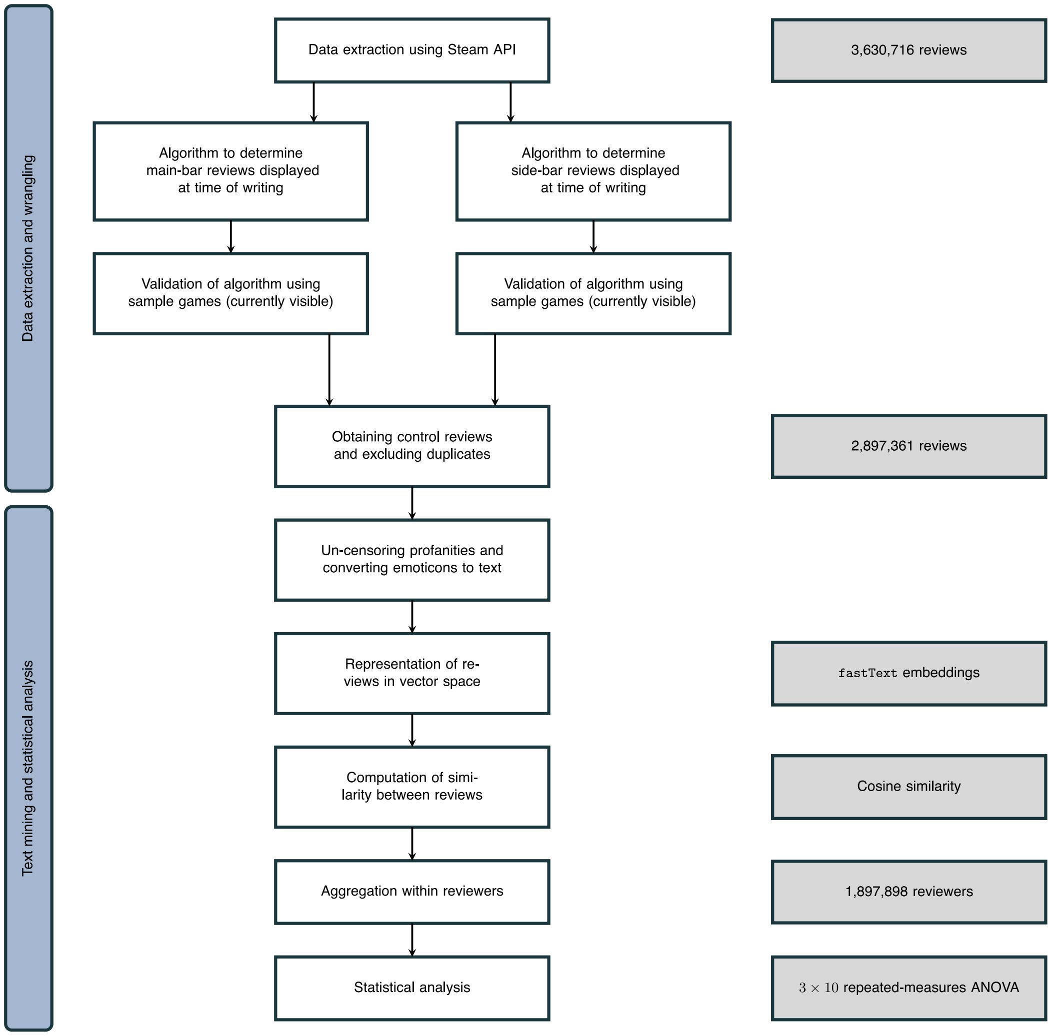
By calculating the similarity of a target review and reviews that were (a) in the main bar, (b) in the sidebar, and (c) *not* visible on the main page (control condition), we can control for idiosyncratic variations in review style across games and time. Importantly, because the control condition includes reviews that were written within the same time window as the reviews displayed in the main bar and the sidebar (i.e., within the 30 days preceding the target review), it controls for a range of time- and game-sensitive features of the successful and recent reviews (e.g., releases of patches, bugs, marketing campaigns, etc.).

To obtain the necessary amount of reviews and compute the variables from those reviews, our study involved the three steps of data extraction and wrangling, text mining, and statistical analyses, which are illustrated in Figure 2 and will be described in detail in the following sections.

Data Extraction and Wrangling

For our study, we collected all reviews from all games on Steam that were tagged as both

Figure 2
A Flowchart Illustration of the Present Study's Methodological Approach and Key Descriptive Statistics



Note. API = application programming interface; ANOVA = analysis of variance. See the online article for the color version of this figure.

"single-player" and "FPS" (first-person shooter), as identified by the "1663%2C4182" tag in the official Steam application programming interface (API). The data were scraped on 26th–27th of January 2024 and resulted in a total of 3,630,716 reviews across 2,304 games. While scraping, some webpages have changed so that 7,117 of the reviews that were scraped were exact duplicates, which we removed. Additionally, the official Steam API did not return any reviews for 1,378 games. A vast majority of these games (944 or 68.5%) indeed had zero reviews, and only a very small proportion (67 or 4.86%) had more than 100 reviews. We did not manually scrape those reviews because our preprocessing pipeline requires an API-exclusive variable (see details below).

In Table 1, we report some descriptive summary statistics of the extracted reviews. One notable feature is the relation between vocabulary size and the number of words. The median vocabulary size (the number of unique words that appear

Table 1
Descriptive Statistics of the Reviews Used in the Analyses

Item	Summary statistic		
	M	*SD*	*Mdn*
Vocabulary size	35.57	64.79	12
Number of words	50.73	117.80	13
Number of reviews per game	1,576	6,291	24
Number of reviews per reviewer	1.53	1.61	1
Time spent playing the game when writing review (in hours)	60.50	297.39	11.90
Upvotes received	5.28	72.54	1

in each review) was only slightly lower than the number of total words, reflecting a low tendency to repeat words and a considerable amount of information contained in the reviews. Additionally, the median playing time before writing a review was 11.90 hr, suggesting that reviewers had some considerable experience with the games before writing a review.

For our subsequent analyses, we needed to identify the reviews that were displayed in the main bar and in the sidebar at the time each of the reviews in our database was composed. We had to reverse-engineer the algorithm that populates the review bars with recent and most helpful reviews because Steam does not provide this information via its API. To reconstruct the sidebar content, where recency-sorted reviews are displayed, we simply obtained the most recent reviews (relative to the time when the target review was written), excluding reviews already shown in the main-bar section of the webpage. As the reviews in the sidebar are ordered according to their recency, we were able to reconstruct the order in which they were displayed to each user. There is no minimum threshold for the number of reviews displayed on the sidebar, but the maximum number of reviews displayed is set to 10. To validate our method for reconstructing the contents of the sidebar, we randomly selected 20 games and correlated the results of our extraction with the true results visible on the website (currently). For ten out of the 20 games, all reviews were present in our validation sample, and only for three of the games, fewer than five reviews were present in the validation sample. The order of reviews that were present in the validation sample was perfectly reconstructed for all but two games.

In the case of the main bar, we attempted to reconstruct its contents using the Steam-provided "weighted_vote_score." Note that the weighted vote score is not the same as the number of helpfulness votes received by each review. This score is accessible through the official API, but it is not directly visible on the Steam webpages. Whenever there are enough reviews to fill the main bar (i.e., at least 10 reviews) that have been composed within the past 30 days and have obtained at least one "helpfulness" vote by other community members, the weighted vote score provides a perfect measure of the order of reviews at the time of display.

If there were not enough reviews published in the last 30 days to fill the main bar, we expanded our research under the assumption that Steam uses an extended window of 30–90 days to select the most helpful reviews. If, at this stage, there are still not enough reviews to be displayed, the same procedure is applied to reviews written in the 90–180 days period, after which all reviews that are older than 180 days are used.

We validated the reconstructed contents of the main bar using the same 20 games that we used to validate the reconstructed contents of the sidebar. For all but two of the 20 games, our algorithm retrieved at least 70% of the games that were displayed in the main bar in real time. The order of reviews that were present in the validation sample was perfectly reconstructed for all but three games.

We removed all games from the analysis for which the algorithm resulted in fewer than 10 reviews that would be displayed in the main or the sidebar (249,259 or 6.9%), after which 3,388,574 reviews from across 868 games remained.

Finally, to determine the control reviews, we identified the 10 most recent reviews written within 30 days before the target review was written and that were neither part of the main bar nor the sidebar. We excluded a small proportion of reviews (123,631; 3.6%) for which there were fewer than 10 reviews written within 30 days in addition to those displayed on the main review page, resulting in a review count of 3,264,943 across 717 games. We further removed 183,791 (6%) of these reviews because they appeared on multiple pages (e.g., when games are parts of different bundles), which resulted in a final review count of 2,897,361 across 694 games.

Text Mining and Statistical Analysis

We preprocessed the reviews for the purpose of subsequent analyses (Kadhim et al., 2014). Steam automatically censors selected profanities in reviews by replacing them with heart (♥) or asterisk (*) symbols. We relied on a community-compiled list of hypothesized censored profanities (Steam Developer Community, 2021) to convert censored words to their original profane text by matching the contents in this list to the length of the censored words. Next, we converted emojis and emoticons back into their original meanings using the emot package for Python. In the next step, we stripped the reviews of stop words (such as "is" or "are") and used lemmatization in combination with grammatical tagging to increase the running speed of subsequent analyses (Loper & Bird, 2002; Brill, 1992; Walkowiak et al., 2018).

To represent the reviews as vectors in a vector space (Bhatia, 2017), we used pretrained `fastText` embeddings (Athiwaratkun et al., 2018; Joulin et al., 2016). We selected `fastText` for two main reasons. First, compared to purely metric-based techniques, `fastText` considers the contextual semantics in addition to the word meanings themselves. Specifically, purely metric-based text embeddings typically assign weights to individual n-grams with minimal consideration of the semantic meaning and relationships between the words (Kasumba & Neumann, 2022; van Tussenbroek, 2020). Second, `fastText` is better able to handle unseen words (Athiwaratkun et al., 2018; Won et al., 2021). `fastText` works by breaking down each word into subword units (also known as vectors of character n-grams). Given that unseen text (e.g., short non-English terms) is expected to occur frequently in video game reviews, `fastText` will still be able to generate a reliable representation by summing up the vectors of its character n-grams.

To compute the similarity between the to-be-written reviews and the vector representations of the visible reviews, we relied on cosine similarity. Formally, cosine similarity $S_{i,j}$ between two vectors i and j is given by

$$S_{i,j} = \frac{V_i \cdot V_j}{\left\|V_i\right\| \cdot \left\|V_j\right\|} = \frac{\sum_{k=1}^{n} (V_i)_k \cdot (V_j)_k}{\sqrt{\sum_{k=1}^{n} (V_i)_k^2} \cdot \sqrt{\sum_{k=1}^{n} (V_j)_k^2}}, \quad (1)$$

where $S_{i,j} \in [-1, 1]$, V is the vector of each review with elements k. Similarity scores of 1 reflect perfect similarity, whereas those with a score of −1 reflect semantically most distant concepts/terms. See Table 2 for an illustration. For each review, we calculated all 30 similarity scores between the target review and the 10 main-bar reviews, the 10 side-bar reviews, and the 10 control reviews according to Equation 1.

For our statistical analysis, we relied on a 3 (salience: main bar vs. sidebar vs. control) by 10 (order: 1–10, from top to bottom) repeated-measures analysis of variance (ANOVA) with the similarity score as the dependent variable. The similarity score was obtained by computing the similarity between the target review and the review at the respective position on the page. For this analysis, we aggregated data within each cell of the experimental design for each reviewer who wrote more than one review. This reduced the number of reviews to 1,897,898. The results do not change qualitatively without aggregation and are provided in the output of the code (Spektor et al., 2024). Due to the large number of observations included in the analysis, we will rely on visual inspection and standardized effect sizes (η_p^2) in addition to formal significance testing.

Results

With respect to our hypotheses, a 3 (salience: main bar vs. sidebar vs. control) by 10 (order: 1 to 10, from top to bottom) repeated-measures ANOVA revealed a main effect of salience, $F(2, 3'795'794) = 169{,}202.66$, $p < .001$, $\eta_p^2 = .082$, a main effect of display order, $F(9, 17'081'073) = 431.50$, $p < .001$, $\eta_p^2 < .001$, and an interaction effect of the two factors, $F(18, 34'162'146) = 395.97$, $p < .001$, $\eta_p^2 < .001$.

In terms of the main effect of salience, post hoc t tests revealed that all groups (main bar, sidebar,

Table 2
Illustration of the Similarity Score and the Effect Sizes Observed in the Study

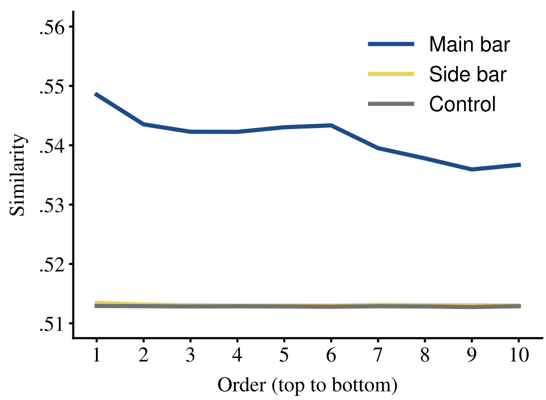
Target review		
"Levels are well designed, offering a great playground for frantic gun battles."		
No.	**Comparison review**	**Similarity score**
1	"Doom's levels are well designed, offering a great playground for frantic gun battles."	.95
2	"Doom is a power fantasy come true, letting you unleash hell on demons in meticulously designed environments."	.58
3	"A visceral and satisfying exploration of pure, unadulterated action, Doom will leave you wanting more."	.51
4	"git gud."	.19

Note. The comparison reviews' similarity scores are displayed with respect to the target review. All reviews were generated artificially to reflect a representative sample of a very similar review (No. 1), two reviews that reflect the average observed similarity scores in the data (Nos. 2 and 3), and one review that is dissimilar to the target review (No. 4). All reviews were preprocessed according to the preprocessing pipeline used for our main analyses.

control) differed significantly from one another. However, the difference between the control and sidebar was negligible (d = 0.001), whereas the main bar had moderately higher values than both the sidebar (d = 0.172) and the control reviews (d = 0.174). The effect sizes are illustrated in Table 2, reviews 2 and 3.

To characterize the interaction, separate one-way repeated-measure ANOVAs were run for each of the three salience conditions separately. The effect of the order was only significant in the main bar, $F(9, 17'081'073) = 1{,}532.19$, $p < .001$, $\eta_p^2 = .001$, but not in the sidebar, $F(9, 17'081'073) = 1.46$, $p = .156$, $\eta_p^2 < .001$, and not for the control reviews, $F(9, 17'081'073) = 0.222$, $p = .992$. The latter of which is noteworthy since this null effect would have been highly unlikely if our algorithms for reconstructing reviews from the three categories had a high rate of misclassifications. After all, the control

Figure 3
Main Results of the Study



Note. Similarity scores reflect the per-reviewer average of the similarity between their reviews and the reviews at the corresponding position on the webpage, as a function of salience (main bar vs. sidebar vs. control) and the display order (from top to bottom). See Figure 1 for an illustration of the webpage format. Error bars are omitted due to the sample size (i.e., the range of the confidence interval is virtually zero). See the online article for the color version of this figure.

reviews were not visible on the main review page, so order should not play a role. The statistical analyses and effect sizes confirm what can be seen in Figure 3: There is a modest effect of salience, such that people's reviews are more similar to reviews displayed in the main bar, and only in the main bar is there a small effect of order, such that reviews are most similar to the reviews displayed at the top of the page.

As can be seen from Table 1, the number of reviews per game is extremely positively skewed. Very few games have many reviews, whereas most games have only very few reviews. Considering that the ANOVAs presented above are representative of the total distribution of reviews, the reviews of games that have many reviews are more frequent in the data set, thus driving the results to a large extent. To rule out that some idiosyncratic properties of those few games with the most reviews create a spurious effect, we investigated the proportion of games in which the observed effects occurred.

This analysis confirmed the main effects of salience: The main-bar similarity score was higher than the side-bar similarity and the control similarity scores for 72.0% and 71.9% of the games, respectively, and the side-bar similarity score was higher than the control similarity score for 55.9% of the games. In the next step, we fit a linear regression with the similarity score as the dependent variable and the order as the only predictor variable for each game and salience condition separately. We found that the slope of order in the main bar was negative for 69.5% of the games and deviated significantly from zero, $t(693) = -6.840$, $p < .001$. The slope of order in the sidebar was negative for 54.0% of the games and did not differ significantly from zero, $t(693) = 0.308$, $p = .758$. The slope of order in the control condition was negative for 52.3% of the games and did not differ significantly from zero, either, $t(693) = 0.514$, $p = .607$.

The results so far suggest the content of a newly written review is most similar to the review that is presented most prominently, namely in the main bar and in the top position, and that is at the same time the review that has been rated as the most helpful by other users. In the following analysis, we asked whether imitating the most successful review at the time is a good strategy for writing a review that itself will become successful. To answer this question, we used linear regression to predict the Z-standardized number of helpfulness votes that a review obtained as a function of the similarity between the top-most review from the main bar and the time stamp at which the review was created (to control for the fact that earlier reviews have had more opportunities to be rated as helpful). This analysis revealed a modest but positive effect of similarity score (standardized $\beta = .019$) and a modest but negative effect of time (standardized $\beta = -.015$). In other words, for each standard deviation above the mean degree of similarity between reviews, the number of helpfulness votes that a review receives increases by 1.9% of its standard deviation.

Discussion

The present study investigated review-writing behavior on a large digital storefront for video games. We hypothesized that customers are influenced by the reviews they see on the screen at the time of composing their own reviews, such that their reviews will be similar to salient reviews that are most easily accessible. Our results supported both of these hypotheses so that successful and salient reviews displayed in the center of the webpage (the "most helpful" reviews published in the last 30 days) had the largest influence on to-be-written reviews. A follow-up analysis showed that reviews that imitate these most helpful reviews tend to be rated as more helpful themselves, reflecting a downstream effect of particularly "successful" reviews.

Our results corroborate past findings highlighting the influence of salience (Buscher et al., 2009; Faraday, 2000; Roth et al., 2013; Shomstein et al., 2019) and ordering (Asad et al., 2021; McCarthy, 2013; Nielsen, 2010) on the attentional allocation of consumers. Our study goes beyond what was previously shown by demonstrating how the increased attentional allocation translates into behavior. One way to interpret the results is that customers writing their reviews do not disregard the information on the screen but rather compose their reviews in a similar fashion to what they see. In other words, to understand and quantify the information content of reviews, it is crucial to consider the context within which it was written. Both positive and negative reviews that were written within a narrow time window could reflect the imitation process studied here rather than the true opinion of the reviewers. If the influence of visible elements was purely a bottom-up effect, then one would expect both the main-bar reviews and side-bar reviews to exert an influence on the review contents. In contrast, there was virtually no difference between the influence of reviews displayed in the sidebar and reviews that were written in the same time frame but that were *not* displayed on the review page. This suggests a directed attempt at imitating successful reviews but ignoring reviews that have not yet received helpfulness ratings. The main effect of order that was only of a somewhat noteworthy size for the main bar reviews further corroborates this interpretation, as the review displayed at the very top of the main bar usually corresponds to the most helpful one.

Although we interpret these results as attempts to mimic successful reviews (and we found that reviews that are more similar to successful reviews tend to be more successful themselves), we acknowledge that we cannot be certain about people's motivation. It is possible that reviewers are selectively influenced by the content of the main bar when composing their own reviews instead of trying to imitate their popularity.

We believe that our approach of aggregation across various features of the reviews and games (e.g., subgenres of games, positivity/negativity of reviews, absolute level of helpfulness, etc.) provides a robust investigation of the qualitative patterns: While we implicitly control for them by comparing the similarity of the visible reviews to reviews that were not directly visible at the time of writing, the observed effect sizes are likely to be attenuated by not explicitly implementing these factors in our analyses. Future studies can apply our methodology to investigate the moderating cognitive and situational factors behind review imitation.

Our results are based on semantic similarity obtained from a large pretrained language model. This measure is suitable for capturing a variety of similarities between two texts in their semantic content. For example, two reviews would be more similar, using our metric, if they both focus on the quality of graphics in a given shooter game rather than if one covers graphical fidelity but the other elaborates on the game's controls. Future research could explore the conditions under which imitation is more prevalent, including factors associated with the products themselves (e.g., life-service shooters vs. single-player games) or other extraneous factors (e.g., times when games receive critical updates). Beyond semantic similarity, imitations could also vary as a function of the review's content (e.g., based on the topic structure of a given review, its length, or its valence [positive vs. negative]).

In sum, the results of our study can be translated into practical implications for commercial game developers, video game enthusiasts, and researchers alike. Particularly salient and helpful reviews influence the to-be-written reviews by increasing the tendency of other reviewers to write similar reviews. This imitation is not without consequences, as these new reviews are more likely to become salient and helpful in the future. Developers may want to try to boost their sales by exploiting the effect. For example, they could ensure that prominently displayed reviews are particularly positive, which could result in a "ripple-like" effect on the spread of positive reviews (Gremler & Brown, 1999). On the flip side, consumers should be careful in inferring the quality of a product based on the most prominent reviews, as they are likely to be, at least partly, imitations of one another. This effect might impair the accumulation of knowledge, and future research should investigate it further.

Data and Code Access Statement

Data used for this study were retrieved using the Steam API. Commercial restrictions on the usage and distribution of Steam's content were abided in this study. The API's terms of use can be found at https://steamcommunity.com/dev/apiterms. Code that scrapes data and runs all analyses reported in the article is publicly available on the Open Science Framework (Spektor et al., 2024). Running the code will naturally yield data that differ from those used in the present study as new games are released, new reviews are written, and existing reviews are edited and/or receive votes from other users. The raw data used in the present study will be provided in case of legitimate interest.

References

Ahmad, S. N., & Laroche, M. (2015). How do expressed emotions affect the helpfulness of a product review? Evidence from reviews using latent semantic analysis. *International Journal of Electronic Commerce*, *20*(1), 76–111. https://doi.org/10.1080/10864415.2016.1061471

Anderson, E. T., & Simester, D. I. (2014). Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research*, *51*(3), 249–269. https://doi.org/10.1509/jmr.13.0209

Asad, M., Halim, Z., Waqas, M., & Tu, S. (2021). An in-ad contents-based viewability prediction framework using artificial intelligence for web ads. *Artificial Intelligence Review*, *54*(7), 5095–5125. https://doi.org/10.1007/s10462-021-10013-3

Ashby, N. J., Walasek, L., & Glöckner, A. (2015). The effect of consumer ratings and attentional allocation on product valuations. *Judgment and Decision Making*, *10*(2), 172–184. https://doi.org/10.1017/S1930297500003934

Athiwaratkun, B., Wilson, A., & Anandkumar, A. (2018). *Probabilistic FastText for multi-sense word embeddings*. [Conference session] Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Melbourne, Australia. https://doi.org/10.18653/v1/P18-1001

Babić Rosario, A., Sotgiu, F., De Valck, K., & Bijmolt, T. H. (2016). The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *Journal of Marketing Research*, *53*(3), 297–318. https://doi.org/10.1509/jmr.14.0380

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the tribes: Using text for marketing insight. *Journal of Marketing*, *84*(1), 1–25. https://doi.org/10.1177/0022242919873106

Bhatia, S. (2017). Associative judgment and vector space semantics. *Psychological Review*, *124*(1), 1–20. https://doi.org/10.1037/rev0000047

Brandes, L., & Dover, Y. (2022). Offline context affects online reviews: The effect of post-consumption weather. *Journal of Consumer Research*, *49*(4), 595–615. https://doi.org/10.1093/jcr/ucac003

Brill, E. (1992). *A simple rule-based part of speech tagger* [Conference session]. Speech and Natural Language: Proceedings of a Workshop, Harriman, New York. https://doi.org/10.3115/974499.974526

Buscher, G., Cutrell, E., & Morris, M. R. (2009). *What do you see when you're surfing? Using eye tracking to predict salient regions of webpages* [Conference session]. SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, United States. https://doi.org/10.1145/1518701.1518705

Cheung, C. M., & Lee, M. K. (2012). What drives consumers to spread electronic word of mouth in online consumer-opinion platforms. *Decision Support Systems*, *53*(1), 218–225. https://doi.org/10.1016/j.dss.2012.01.015

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. https://doi.org/10.1509/jmkr.43.3.345

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222. https://doi.org/10.1146/annurev.ne.18.030195.001205

Eberhard, L., Kasper, P., Koncar, P., & Gutl, C. (2018). *Investigating helpfulness of video game reviews on the Steam platform* [Conference session]. 5th International Conference on Social Networks Analysis, Management and Security, Valencia, Spain. https://doi.org/10.1109/SNAMS.2018.8554542

Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, *52*, 498–506. https://doi.org/10.1016/j.tourman.2015.07.018

Faraday, P. (2000). Visually critiquing webpages. In N. Correia, T. Chambel, & G. Davenport (Eds.), *Multimedia'99. Eurographics* (pp. 155–166). Springer. https://doi.org/10.1007/978-3-7091-6771-7_17

Gremler, D. D., & Brown, S. W. (1999). The loyalty ripple effect: Appreciating the full value of customers. *International Journal of Service Industry Management*, *10*(3), 271–293. https://doi.org/10.1108/09564239910276872

Grewal, L., & Stephen, A. T. (2019). In mobile we trust: The effects of mobile versus nonmobile reviews on consumer purchase intentions. *Journal of Marketing Research*, *56*(5), 791–808. https://doi.org/10.1177/0022243719834514

Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., & Mikolov, T. (2016). *Fasttext.zip: Compressing text classification models*. arXiv. https://doi.org/10.48550/arXiv.1612.03651

Kadhim, A. I., Cheah, Y. N., & Ahamed, N. H. (2014, December). *Text document preprocessing and dimension reduction techniques for text document clustering* [Conference session]. In 2014 4th international conference on artificial intelligence with applications in engineering and technology, Kota Kinabalu, Malaysia. https://doi.org/10.1109/ICAIET.2014.21

Kapoor, G., & Piramuthu, S. (2009). Sequential bias in online product reviews. *Journal of Organizational Computing and Electronic Commerce*, *19*(2), 85–95. https://doi.org/10.1080/10919390902820764

Kasumba, R., & Neumann, M. (2022). *Practical sentiment analysis for education: The power of student crowdsourcing* [Conference session]. Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, Canada. https://doi.org/10.1609/aaai.v38i21.30356

Kosslyn, S. M., & Alper, S. N. (1977). On the pictorial properties of visual images: Effects of image size on memory for words. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *31*(1), 32–40. https://doi.org/10.1037/h0081649

Kronrod, A., & Danziger, S. (2013). "Wii Will Rock You!" The use and effect of figurative language in consumer reviews of hedonic and utilitarian consumption. *Journal of Consumer Research*, *40*(4), 726–739. https://doi.org/10.1086/671998

Lafky, J. (2014). Why do people rate? Theory and evidence on online ratings. *Games and Economic Behavior*, *87*, 554–570. https://doi.org/10.1016/j.geb.2014.02.008

Liu, X., Lee, D., & Srinivasan, K. (2019). Large-scale cross-category analysis of consumer review content on sales conversion leveraging deep learning. *Journal of Marketing Research*, *56*(6), 918–943. https://doi.org/10.1177/0022243719866690

Loper, E., & Bird, S. (2002). *Nltk: The natural language toolkit* [Conference session]. ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics, Philadelphia, PA, United States. https://doi.org/10.3115/1118108.1118117

McCarthy, C. (2013). *To scroll or not to scroll: The age-old question* [Conference session]. Proceedings Computing and Information Technology Research and Education, New Zeland. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=e3041ca454f173de7893aa9b5c26e929152c4209

Moore, S. G. (2015). Attitude predictability and helpfulness in online reviews: The role of explained actions and reactions. *Journal of Consumer Research*, *42*(1), 30–44. https://doi.org/10.1093/jcr/ucv003

Naylor, R. W., Lamberton, C. P., & Norton, D. A. (2011). Seeing ourselves in others: Reviewer ambiguity, egocentric anchoring, and persuasion. *Journal of Marketing Research*, *48*(3), 617–631. https://doi.org/10.1509/jmkr.48.3.617

Nielsen, J. (2010). *Scrolling and attention*. Nielsen Norman Group. https://www.nngroup.com/articles/scrolling-and-attention-original-research/

Offerman, T., & Sonnemans, J. (1998). Learning by experience and learning by imitating successful others. *Journal of Economic Behavior & Organization*, *34*(4), 559–575. https://doi.org/10.1016/S0167-2681(97)00109-1

Roth, S. P., Tuch, A. N., Mekler, E. D., Bargas-Avila, J. A., & Opwis, K. (2013). Location matters, especially for non-salient features—An eye-tracking study on the effects of web object placement on different types of websites. *International Journal of Human–Computer Studies*, *71*(3), 228–235. https://doi.org/10.1016/j.ijhcs.2012.09.001

Schindler, R. M., & Bickart, B. (2012). Perceived helpfulness of online consumer reviews: The role of message content and style. *Journal of Consumer Behaviour*, *11*(3), 234–243. https://doi.org/10.1002/cb.1372

Schoenmueller, V., Netzer, O., & Stahl, F. (2020). The polarity of online reviews: Prevalence, drivers and implications. *Journal of Marketing Research*, *57*(5), 853–877. https://doi.org/10.1177/0022243720941832

Shomstein, S., Malcolm, G. L., & Nah, J. C. (2019). Intrusive effects of task-irrelevant information on visual selective attention: Semantics and size. *Current Opinion in Psychology*, *29*, 153–159. https://doi.org/10.1016/j.copsyc.2019.02.008

Spektor, M. S., Walasek, L., & Alba, C. (2024). *Code for "Attention-driven imitation in consumer reviews."* https://doi.org/10.17605/OSF.IO/73YVJ

Steam Developer Community. (2021). *English-profanity.txt*. GitHub. https://github.com/kast1450/steam-profanity-filter/blob/main/english/english-profanity.txt

van Laer, T., Edson Escalas, J., Ludwig, S., & van den Hende, E. A. (2018). What happens in vegas stays on tripadvisor? A theory and technique to understand narrativity in consumer reviews. *Journal of Consumer Research*, *46*(2), 267–285. https://doi.org/10.1093/jcr/ucy067

van Tussenbroek, T. (2020). *Who said that? Comparing performance of TF-IDF and fastText to identify authorship of short sentences* [Master's thesis, TU Delft]. https://resolver.tudelft.nl/uuid:93873bbf-2886-4023-b696-e11be2b99024

Vana, P., & Lambrecht, A. (2021). The effect of individual online reviews on purchase likelihood. *Marketing Science*, *40*(4), 708–730. https://doi.org/10.1287/mksc.2020.1278

Walkowiak, T., Datko, S., & Maciejewski, H. (2018). *Reduction of dimensionality of feature vectors in subject classification of text documents* [Conference session]. 18th International Conference on Reliability and Statistics in Transportation and Communication, RelStat'18, Riga, Latvia. https://doi.org/10.1007/978-3-030-12450-2_15

Wang, C., Kalra, A., Borcea, C., & Chen, Y. (2015). *Viewability prediction for online display ads* [Conference session]. 24th ACM International on Conference on Information and Knowledge Management, Melbourne, Australia. https://doi.org/10.1145/2806416.2806536

Won, M.-S., Choi, Y., Kim, S., Na, C., & Lee, J.-H. (2021). *An embedding method for unseen words considering contextual information and morphological information* [Conference session]. 36th Annual ACM Symposium on Applied Computing, Virtual Event, Republic of Korea. https://doi.org/10.1145/3412841.3441982

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(5), 601–621. https://doi.org/10.1037/0096-1523.10.5.601

Yoo, K. H., & Gretzel, U. (2008). What motivates consumers to write online travel reviews? *Information Technology & Tourism*, *10*(4), 283–295. https://doi.org/10.3727/109830508788403114

Received June 18, 2023
Revision received March 4, 2024
Accepted March 13, 2024 ■