Design of review systems – A strategic instrument to shape online reviewing behavior and economic outcomes

Dominik Gutt[a],[*], Jürgen Neumann[a], Steffen Zimmermann[b], Dennis Kundisch[a], Jianqing Chen[c]

[a] Paderborn University, Warburger Str. 100, 33098 Paderborn, Germany
[b] Universität Innsbruck, Innrain 52, 6020 Innsbruck, Austria
[c] The University of Texas at Dallas, Jindal School of Management, 800 West Campbell Road, SM 33, Richardson, TX 75080, United States

ARTICLE INFO

ABSTRACT

As online reviews play a decisive role in consumers' purchase decisions, e-commerce platforms are using review systems strategically to obtain a competitive advantage. However, the strategic potential can only be leveraged if the review system is designed appropriately. Research on the design of review systems and the effects of design choices has not yet been summarized or synthesized in a review article. We aim to close this gap by providing a scoping review. In our synthesis we posit that the design of review systems moderates the impact of online reviews on economic outcomes and the factors that drive the formation of reviews. After reviewing current research findings, we identify gaps and provide a research agenda covering three key themes: *Design features*, *environments*, and *devices*.

Introduction

Online reviews provide firms with strategic knowledge that is pivotal for price setting, demand forecasting, product quality assessment, and customer relationship management. The full exploitation of the strategic potential of online reviews, however, is contingent on the underlying IT system. Bharadwaj (2000) notes that IT systems "such as groupware and expert systems, when populated with firm-specific knowledge and insights, are transformed into specialized assets that are almost impossible to imitate by competitors (p. 175)". Following Bharadwaj we argue similarly that review systems, when populated with online reviews, represent specialized assets in the form of the (reviewing) consumers' experiences and knowledge. According to the resource-based theory of the firm (Barney, 1991), review systems thus fulfill the necessary condition of representing a *valuable*, *rare*, *inimitable*, and *non-substitutable* resource to the firm, with which it can obtain a competitive advantage.[1] Consequently, review systems have evolved into strategic information systems that are used for that very purpose by e-commerce platforms such as Amazon or Yelp (Gable, 2010). The strategic importance of developing review systems into strategic information systems by making the "*right*" design choices is exemplified by Amazon's feature which asks consumers to answer the question, "Was this review helpful?". According to Spool (2009), this simple feature alone is responsible for more than US $2.7 billion of new revenue for Amazon every year.

Previous literature reviews of online reviews have started synthesizing the current state of knowledge and presented research

---

* Corresponding author.
  E-mail addresses: dominik.gutt@wiwi.uni-paderborn.de (D. Gutt), juergen.neumann@wiwi.uni-paderborn.de (J. Neumann),
Steffen.Zimmermann@uibk.ac.at (S. Zimmermann), dennis.kundisch@wiwi.uni-paderborn.de (D. Kundisch), chenjq@utdallas.edu (J. Chen).

[1] In Table A1 in the appendix, we contrast the definitions of these four properties with regard to their applicability to review systems as specialized assets.
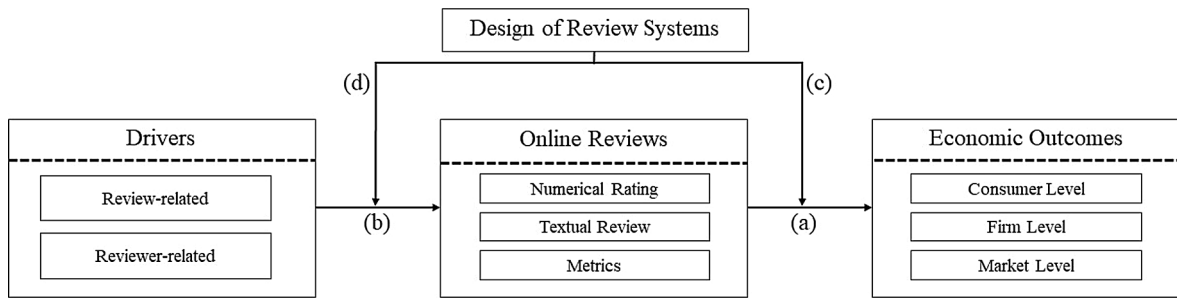
Fig. 1. Conceptual model of existing research.

findings regarding two aspects: (1) the impact of online reviews on economic outcomes – which we refer to as *direct outcome effect* in the following – such as prices and sales (Cheung and Thadani, 2012; Floyd et al., 2014; King et al., 2014; You et al., 2015; Babic Rosario et al., 2016), and (2), the factors that drive online review generation – which we refer to as *direct driver effect* in the following – such as reviewing motivation or reviewer self-selection (Matos and Rossi, 2008; King et al., 2014; Hong et al., 2017). Yet, to the best of our knowledge, a literature review consolidating the growing body of literature on the design of review systems is missing to date. Thus, we undertake a review, guided by the following three research questions:

1. *What is the state-of-the-art in review system design?*
2. *What are the remaining research gaps in review system design?*
3. *What are promising ways to close the research gaps?*

To answer these research questions, we perform a scoping review (Paré et al., 2015). In total, we identify 312 studies on online reviews, originating in diverse disciplines such as information systems, marketing, and strategy. Of these, 58 studies explicitly analyze the *moderating effect* of review system design (*direct driver effect* and *direct outcome effect*). From this literature we derive three research gaps, which we structure around three key themes: (1) *Design features*: A substantial number of studies investigate the economic outcomes or the drivers of online reviews and, at the same time, propose implications for the design of review systems, such as providing reviewers with a predefined review template (e.g., Yin et al., 2014). Although a considerable number of review system design features have been proposed over the years, only very few have actually been analyzed. (2) *Environments:* In recent years, new online business models and environments have emerged including two-sided platform businesses (e.g., Airbnb, Uber). These enable two-sided reviews and require adapted design features to, for instance, mitigate reciprocity in two-sided review systems. (3) *Devices*: Most design features of review systems have been analyzed for stationary devices such as personal computers. However, online reviews are increasingly produced and consumed via mobile devices, which require specific design features. We propose three research thrusts where we illustrate how these gaps can be closed.

Background and research model

The first generation of e-commerce platforms, such as Amazon and eBay, enabled geographically remote buyers to place orders online. To establish trust with the sellers and facilitate transactions, eBay and Amazon installed review systems. Since then a wide range of e-commerce platforms, including Yelp, Airbnb, or TripAdvisor, have embraced review systems, which have become an integral feature of online purchases.

Soon after their inception, the phenomenon of review systems has sparked the interest of scholars who have since then produced a large body of research on online reviews. Fig. 1 represents the conceptual model we use to classify research on online reviews. *Online reviews* usually consist of at least two components, namely a *numerical rating* (e.g., a star rating) and a *textual review*. The numerical rating represents the reviewer's assessment of a product or service while the textual review element complements the numerical rating with additional information. Review systems typically provide various *metrics* that evaluate or aggregate online reviews. Such metrics include, for instance, individual-level metrics, like the perceived *helpfulness* of an online review, and aggregate-level metrics, like the *volume* (i.e., number of online reviews), the *valence* (i.e., the average numerical rating), and the *variance* (i.e., the numerical rating distribution).

Economic *outcomes* of online reviews typically represent the effects directly resulting from the metrics of online reviews (i.e., direct outcome effect (a) in Fig. 1) and can be measured on the *consumer*, the *firm*, or the *market level*. For example, an increasing valence of online reviews may increase a product's sales and thus represent a firm-level direct outcome effect. *Drivers*[2] refers to any effects that influence individual online reviews or any online review metric (i.e., direct driver effect (b)) and, in turn, can be *review-related* or *reviewer-related*. For example, the social influence bias is a review-related driver which suggests that reviewers change their own reviewing behavior when exposed to existing reviews (Muchnik et al., 2013).

---

[2] We recognize that earlier literature often refers to biases instead of drivers. Because the term "bias" implies a divergence from an objectively correct statement and such correctness is difficult to establish, we propose the term "driver" instead.

A more recent stream of research investigates the design of review systems as a moderator of the direct outcome effect (i.e., moderating outcome effect (c)) or the direct driver effect (i.e., moderating driver effect (d)). For instance, the cardinality of the rating scale (e.g., binary versus 1 to 5 stars) moderates the relationship between valence and sales, with a high rating scale cardinality seen as benefiting the sale of mainstream products (Jiang and Guo, 2015). Regarding the moderating driver effect, some system designers allow reviewers to provide identity-descriptive information (such as a real name, for example), which has the effect of increasing the perceived helpfulness of these reviews (Forman et al., 2008). One core aim of our research is to identify all the studies that analyze the moderating effect of such design features.

Literature search methodology

In step 1, all authors jointly determined the set of relevant outlets for our study. We decided to include only high-quality journals in order to synthesize established and peer-reviewed scientific knowledge. The inclusion criteria were as follows: (a) starting with previous review studies, we included all the journals these reviews had searched (Cheung and Thadani, 2012; You et al., 2015; Floyd et al., 2014); (b) as journals in the information systems discipline are predestined to publish studies on the *design* of review systems, we included all journals of the AIS Senior Scholars' Basket of Journals, supplemented by the design science-oriented outlet *Business & Information Systems Engineering* and by *Information & Management*; (c) as online reviews are also covered in other disciplines, we enriched the list by journals used in the Financial Times Research Rank 2016 in the area of management, strategy, operations, economics, and marketing. The final list contains 38 journals and can be found in Table A2 in the appendix.[3]

In step 2, each journal was assigned to one of the co-authors. To ensure that we capture all the relevant research articles from our focal journals, we conducted a manual issue-by-issue search between 1991 and 2017 (including "online first" articles) for relevant articles, based on the title and abstract. For each article the co-author decided whether the article deals with some or other aspect of online reviews.

In step 3, a different co-author performed a thematic analysis inspired by Roberts et al. (2012) using our conceptual model (see Fig. 1) to classify each relevant study. Any articles which could not be classified as addressing any of the relationships in our model (arrows (a) to (d) in Fig. 1) were discarded. For the 312 articles that could be classified, a brief coding following a standardized template was performed for the articles covering direct effects (254 articles), whereas a comprehensive coding was performed for the articles analyzing moderating effects (58 articles).[4] The brief coding included a classification according to the direct effects (direct outcome effect or direct driver effect) and the main findings. The comprehensive coding additionally included, amongst others, a classification according to the moderating effects (moderating outcome effect or moderating driver effect), the analyzed design features, information about the independent and dependent variables, the research method, and a characterization of the data.[5]

We used four measures to ensure a consistent quality for our procedure:

- For step 2, we performed an additional keyword-based search[6] for a subsample of three non-IS journals, to avoid any relevant articles being missed out accidentally. Reassuringly, this robustness check did not produce any additional articles.
- For step 2, interrater coding between two co-authors, and between graduate student assistants and co-authors was conducted on a subsample of nine journals (Cohen's Kappa and Krippendorf's Alpha between 0.73 and 1).
- Before conducting step 3, the co-authors jointly used twelve sample articles identified as relevant in step 2 to establish a common understanding of the inclusion and categorization criteria in step 3, and of the subsequent coding procedure.
- After coding all articles, papers classified as design-related articles were discussed in detail by the group of co-authors to ensure correct classification and coding for the 58 articles at the center of our study. For the very few instances of disagreement, the co-authors discussed their opinions to come to a joint verdict about inclusion or exclusion, as well as the coding (Paré et al., 2015).

Synthesis of research findings and research gaps

To guide the reader on the terminology we used in the following discussion, we list the design feature categories synthesized from the relevant literature (see Section 'Literature search methodology') and provide a description of each in Table 1.

We structure this section by distinguishing between articles addressing outcome effects and those addressing driver effects. In each subsection, we start by briefly describing the direct effects before focusing on the moderating effects of design features with reference (in squared brackets) to the design feature categories listed in Table 1. We conclude this section with a discussion of the identified research gaps.

Outcome effects

Consumer level

On the consumer level, studies have investigated direct outcome effects of the informational value of online reviews on consumer

---

[3] The appendix is available online at https://go.upb.de/jsis2019appendix.

[4] A complete list of all 58 studies on moderating effects can be found in Table A3 in the appendix.

[5] The coding template and a graphical illustration of the search strategy can be found in Table A4 and Fig. A1 in the appendix.

[6] For a list of used search terms, see Table A5 in the appendix.

Table 1
Design categories.

Design feature category	Description
Review templates	Provide reviewers with templates and guidelines to improve the reviews they produce
Review presentations	Influence the order in which reviews are presented; provide ranking and filter options of reviews for review readers; change the appearance of individual reviews (e.g., introducing information on the reviewer)
Adapted metrics	Adapt existing metrics like valence, volume, or variance
Rating dimensions	Represent the numerical rating of a review (e.g., multi- vs. single-dimensional, binary vs. 1–5 rating scale)
Management responses	Enables sellers to publicly reply to online reviews
Review elicitation	Asks past customers to provide a review; provide non-monetary and monetary incentives for writing reviews
Reviewer reputation	Introduces concepts to describe the status of a reviewer in the reviewer community (e.g., ranking of reviewers, friendships between reviewers)
Mutual reviewing	Enables both buyers and sellers to review each other, and adjustments to be made to the mutual reviewing process (e.g., double-blind reviewing)
Recommender systems	Introduce a system that provides consumers with product or service recommendations, adjust recommender systems based on their relationship with online reviews
Dishonest behavior	Detects and mitigates fraudulent reviewer and seller behavior (e.g., fake review filters, verified review mechanisms, measures to punish dishonest sellers)

outcomes such as learning (Dellarocas, 2003; Hu et al., 2017a; Wu et al., 2015; Koh et al., 2010), consumer satisfaction (Benlian et al., 2012), and disposition towards the product (Ein-Gar et al., 2012). For example, if platforms feature online reviews or recommendations, consumer satisfaction (and related constructs like perceived ease of use) are higher because the platform facilitates the consumer search process (Benlian et al., 2012).

Moderating outcome effects

First, the review system design can support a consumer's learning process by presenting reviews with the type of information that is most relevant to each stage of the purchasing process (Li et al., 2017; Huang et al., 2014) [Review Presentation]. Allowing videos in review texts can further enhance consumer learning as compared to simple images or text-only reviews, as consumers evaluate these reviews as being more credible and persuasive (Xu et al., 2015).

Second, system design can help increase the perceived ease of use and satisfaction with the e-commerce platform. This includes adapting the rating dimensions to the features of the traded products (Fang et al., 2014) [Rating Dimensions], or using automatic recommendations (Benlian et al., 2012; Hostler et al., 2011) [Recommender Systems].

Third, when consumers are presented with a list of both positive and negative reviews, it can enhance their disposition towards a product (Ein-Gar et al., 2012) [Review Presentation]. The same effect means that, when consumers are faced with highly disaggregated online ratings (i.e., lacking aggregated metrics like average rating or variance), it benefits sellers of products with negative outlier reviews (Camilleri, 2017) [Adapted Metrics].

Firm level

A number of studies investigate the direct effect of online reviews on sales. Evidence from observational data and field experiments supports the hypothesis that metrics – such as *helpfulness*, *valence* or *volume* – of online reviews can causally lead to higher sales (e.g., Chevalier and Mayzlin, 2006, Duan et al., 2008, Forman et al., 2008). Some studies show that implementing product/service *recommendations* also increases sales, due to facilitating consumer search processes (Cheung et al., 2003). Scholars have also examined the direct outcome effect on price (e.g., Ba and Pavlou, 2002; Pavlou and Dimoka, 2006) finding a positive relationship between the valence of a seller's online ratings and her pricing power.

Moderating outcome effects

Regarding the *valence* of online reviews, system designers can influence the representation of the rating by manipulating the scale and dimensionality of ratings [Rating Dimensions]. One study finds that using low rating scale cardinality for niche products and a high rating scale cardinality for mainstream products both benefits sales (Jiang and Guo, 2015).

By positioning reviews and ratings, the influence of review *valence* can also be manipulated [Review Presentation]. Displaying average ratings in the product list may prevent consumers from discovering niche products that actually match their needs while they enhance the sales of main stream products (Li, 2017). When the review system allows sellers to embed their online reviews into their product description, they can increase their sales using this feature (Wang et al., 2016).

Review system design can moderate the impact that *recommendations* can have on sales [Recommender Systems]. Empirical evidence indicates that when displaying recommendations for competing products, the rating valence of the latter decreases the sales of the focal product (Jabr and Zheng, 2014). Balancing the individual relevance of and the profit derived from recommendations has a positive effect on a firm's profit without negatively affecting consumer trust (Panniello et al., 2016). Review texts (Ghose et al., 2012), review sequence (Piramuthu et al., 2012), and missing ratings (Ying et al., 2006) can be effectively used to improve the performance of recommender algorithms.

Market level

A central aspect of market efficiency is that all market participants *honestly participate in transactions* so that participants can assess the quality of offers from the *valence* and the *volume* of reviews. However, due to the anonymity of buyers and sellers in electronic markets, the risk for moral hazard can be high, especially for sellers (Dellarocas, 2003).

Moderating outcome effects

One study finds that review systems improve the performance of consumer-to-consumer (C2C) auction markets and that punishing cheating market participants is more effective than rewarding honest participants (Yang et al., 2007) [Dishonest Behavior]. Design features such as the granularity of feedback [Review Dimensions], the format of the reputation profile, or the policy regarding missing feedback [Review Presentation], can effectively mitigate moral hazard and, in turn, increase market efficiency (Dellarocas, 2005).

Regarding the role of review valence for market efficiency, studies have demonstrated that updating a seller's online review profile only every *k* transactions instead of after every transaction can increase the market efficiency (Dellarocas, 2006). Flexible time windows, depending on the seller type, can improve market efficiency even further (Aperjis and Johari, 2010) [Adapted Metrics]. Moreover, in a market where sellers offer heterogeneous goods, the valence is closer to the actual quality if instead of a single rating score for the seller there is one for each good offered (Samak, 2013) [Adapted Metrics]. Reciprocity between buyers and sellers, whereby they mutually rate each other following a purchase, can lend an upward bias to the online ratings [Mutual Reviewing]. If all the ratings are high irrespective of the underlying purchase, online ratings are "inflated", and hence fail to discriminate between supposedly "good" and "bad" trading partners (Bolton et al., 2013). Thus, reciprocity in C2C markets should be accounted for by the review system design, e.g., by only allowing for blind, one-sided and anonymous ratings, in order to avoid inflation and promote market efficiency (Bolton et al., 2013).

Studies have also provided empirical evidence that sellers (Li and Xiao, 2014) and platform owners (Avery et al., 1999) can increase the volume of reviews by granting rebates to consumers who provide an online rating [Review Elicitation]. Even though feedback obtained in this way is mostly upward-biased, it can increase the market efficiency as buyers are more likely to provide an online rating (Li and Xiao, 2014).

Driver effects

Naturally, design choices are an important strategic instrument to achieving desirable economic outcomes. These outcomes often derived from positive reviews (e.g., Chevalier and Mayzlin, 2006), a high number of reviews (e.g., Duan et al., 2008), or a high number of helpful reviews (e.g., Forman et al., 2008). It is important therefore to discuss research that considers design features which moderate direct driver effects.

Helpfulness

First, *trust in the reviewer* directly affects perceived helpfulness. If the author of a review is not perceived as credible, reviews are not perceived as helpful (Schlosser, 2011; Chen and Lurie, 2013). Second, there exists a relationship between *trust in the review* and perceived helpfulness. Trust in the review is usually based on information provided in the review text and the rating. For instance, if the rating strongly deviates from the average rating, a review is less credible and thus deemed less helpful (Yin et al., 2016). Third, the reviewer's *motivation* to write the review plays a role in how helpful the review will become. If reviewers are sufficiently motivated to spend time and effort to write a review, the information they convey can be more helpful (Korfiatis et al., 2012).

Moderating driver effects

By letting reviewers add identity-descriptive information (e.g., name, geographic location or profile picture) to their reviews, system designers can support reviewers in improving their trustworthiness and thus increase the perceived helpfulness of their reviews (Forman et al., 2008; Karimi and Wang, 2017) [Review Presentation]. Another way of increasing the perceived helpfulness of reviews involves awarding badges or certificates to reputable reviewers (Kuan et al., 2015; Chang and Wu, 2014; Chang et al., 2013) [Reviewer Reputation].

Systems can support reviewers in writing *trustworthy reviews* [Review Presentation]. If a system allows reviewers to add images or videos to their reviews, perceptions of helpfulness increase depending on the product type (Xu et al., 2015). To further increase the trust in reviews, systems can employ user-controllable filters to select reviews based on different categories (e.g., filter reviews on TripAdvisor by travel season) (Hu et al., 2017b) [Review Presentation].

The effect of reviewer *motivation* and helpfulness can be negatively moderated by implementing a design feature that allows asking consumers to become a reviewer [Review Elicitation]. It has been found that asking past consumers to write a review by email leads to reviews that are perceived as less helpful than reviews that consumers have provided spontaneously (Askalidis et al., 2017).

Volume

Several studies suggest that reviewers have several intrinsic sources of *motivation* to write a review (e.g., Hennig-Thurau et al., 2004). For instance, consumers write reviews to help the company or to retaliate after a negative experience. However, writing a review is costly in terms of time and effort. Since a consumer's motivation rarely outweighs the cost of writing a review, they refrain from publishing one, leading to an underreporting bias (Hu et al., 2017a).

Moderating driver effects

Design features can moderate the direct effect of *motivation* on volume of reviews by providing external non-monetary or monetary incentives [Review Elicitation]. Review system designers can implement social comparisons [Review Elicitation] (Chen et al., 2010) or management responses [Management Responses] (Proserpio and Zervas, 2015) as a new feature to positively moderate a reviewer's motivation and increase the review volume. Regarding monetary incentives, enabling sellers to give a rebate to consumers might make consumers more likely to write a review (Li and Xiao, 2014; Chen et al., 2017).

The relationship between the reviewer's *motivation* and the volume of reviews is also influenced by the desire to gain reputation [Reviewer Reputation]. If the review system design allows for follower or friendship relationships, reviewers with a high number of relationships will publish more reviews than those with a low number (Sun et al., 2017; Goes et al., 2014). Thus, publicly available reputation information encourages consumers to write more reviews. Nevertheless, such reviewer reputation mechanisms can also entail disadvantages because highly reputable reviewers tend to refrain from reviewing popular products (Shen et al., 2015).

Finally, double blind reciprocal reviewing, which is common practice on two-sided platform like Airbnb, can decrease the propensity to write a review (Bolton et al., 2013) [Mutual Reviewing].

Valence

Naturally, reviewers give higher ratings if they achieve a higher *taste match* with the product or service (Sun, 2012). Moreover, as consumers with a higher preference for a product are more likely to buy and review a product, online reviews tend to be more positive. This effect is called *preference bias* (Li and Hitt, 2008; Hu et al., 2017a). Due to the *social influence bias*, reviewers change their own assessment of a product/service based on seeing existing reviews (Muchnik et al., 2013). If previous ratings are positive, reviewers tend to change their intended rating upwards. If reviews are negative, but the reviewer's own experience is positive, she may aim to correct the previously existing negative ratings by giving an even higher rating. The *price bias* in single-dimensional rating systems leads to a decrease in ratings when the price for the product increases because of the absence of a "value for price" dimension (Li and Hitt, 2010).

Another driver of review valence is *attention.* If reviewers expect attention from other consumers or the seller, they try to be more objective in their texts and ratings (Shen et al., 2015; Proserpio and Zervas, 2015). Reviewers also change their rating behavior if their rating could influence their *own reputation*. On a mutual reviewing platform (e.g. AirBnB), both sides consider any unwanted effects on their own reputation when providing a review (Dellarocas and Wood, 2008). Some reviewers also try to weaken the online review score of businesses by giving fake reviews or trying to enhance the review score by giving fake positive reviews (e.g., Mayzlin et al., 2014). In this case, the valence of online reviews is clearly driven by *dishonest intentions*.

Moderating driver effects

Compared with single-dimensional rating systems, multi-dimensional rating systems make it easier for consumers to find products or services that match their preferences. This design feature, then, positively moderates the *taste match* effect (Chen et al., 2018) [Rating Dimensions]. To reduce the impact of previous ratings and address the *social influence bias*, designers can design their system so that it emails consumers asking them to write reviews on a page where they are not exposed to existing reviews (Askalidis et al., 2017) [Review Elicitation]. Allowing friendships between reviewers on the system can also positively moderate the strength of the social influence bias [Reviewer Reputation]. Enabling sellers to offer rebates to buyers interacts with the *price bias* and implies that reviews tend to be more positive (Li and Xiao, 2014) [Review Elicitation]. Employing a multi-dimensional rating scale alleviates this bias (Li and Hitt, 2010) [Rating Dimensions].

Introducing a reviewer reputation system encourages more differentiated reviews because reviewers become aware of the *attention* they receive (e.g., Shen et al., 2015) [Reviewer Reputation]. Moreover, reviewers could also receive attention from sellers. Enabling management responses yields more positive reviews as sellers who respond to negative reviews receive more positive reviews in the future (Proserpio and Zervas, 2015; van Noort and Willemsen, 2012) [Management Responses]. If, however, sellers only respond to some negative reviews, consumers who do not receive a response become less satisfied in the future and give lower ratings (Gu and Ye, 2014).

Systems that allow mutual reviewing exhibit design features that moderate the direct driver effect of a reviewer's *own reputation* [Mutual Reviewing]. Fearing retaliation from the other party, both buyers and sellers refrain from providing negative reviews (Bolton et al., 2013). Similarly, removing the seller's ability to retaliate against negative reviews from buyers increases the number of negative reviews (Ye et al., 2014). This transition to a one-sided review system prevents low-quality sellers from disguising as high-quality sellers and encourages them to improve the quality of their products (Ye et al., 2014). Furthermore, the design feature by which a negative review can be withdrawn after a dispute resolution encourages both parties to provide initially more negative reviews to improve each other's bargaining position (Bolton et al., 2018). Furthermore, sellers often exit and re-enter the system to remove their negative reputation. Assigning new sellers a minimum level of reputation offers a possible remedy (Zacharia et al., 2000) [Adapted Metrics].

To moderate the effects of *dishonest intentions* and reduce their effect on the valence of online ratings, review systems can implement a verified reviewing policy [Dishonest Behavior]. Under this policy, only consumers with a verified purchase are allowed to give a rating, which substantially mitigates the problem of fraudulent reviews (Mayzlin et al., 2014). Additionally, review system designers can develop algorithms to detect fake reviews by considering the sentiment of a review text (Hu et al., 2012), linguistic styles (Ong et al., 2014) or reviewer characteristics (Chung et al., 2013).

To summarize this section, Table 2 depicts the level of outcome effects (market, firm, or consumer) and driver effects (helpfulness, volume, or valence) for each of the moderating effect of the design feature categories investigated. It illustrates existing connections

Table 2

Design feature categories and moderating effects on different levels.

Design feature category	Review → Outcome (Section 'Outcome effects')			Driver → Review (Section 'Driver effects')		
	Consumer	Firm	Market	Helpfulness	Volume	Valence
Review templates						
Review presentations	X	X	X	X		
Adapted metrics	X		X			X
Rating dimensions	X	X				X
Management responses					X	X
Review elicitation			X	X	X	X
Reviewer reputation				X	X	X
Mutual reviewing			X		X	X
Recommender systems	X	X				
Dishonest behavior			X			X

between design feature categories and online review effect levels.

Research gaps

While synthesizing the research findings, we noticed that there are many design features that have been proposed but not yet analyzed, especially in articles that focus on direct effects. Moreover, the majority of studies analyze design features only in a classical Business-to-Consumer (B2C) environment (such as Amazon or TripAdvisor), and only for stationary devices (such as desktop computers or notebooks). However, some studies indicate that different environments (Mayzlin et al., 2014; Neumann and Gutt, 2017) and devices used for posting or reading reviews (Burtch and Hong, 2014) result in differences in moderating effects. Hence we extended our coding accordingly (see Table A2 in the appendix):

1. *Design Features*: Have design features been proposed? (coded for all 312 articles, see Section 'Literature search methodology')
2. *Environments*: In which environments have the design features been analyzed? (coded for the 58 articles covering moderating effects)
3. *Devices*: For which devices have the design features been analyzed? (coded for the 58 articles covering moderating effects)

Design features

Arguably, the most intuitive direction for future research is to analyze design features which have been proposed in previous research but not yet analyzed. A further direction could also be to identify and analyze new design features that were not researched so far. Thus, we formulate our first research gap:

Research Gap 1: Multiple design features for review systems have not yet been analyzed.

Proposed design features. In Table 3 we illustrate selected proposed but yet not analyzed design features (center column) based on the extended coding procedure. In particular, the design feature categories "Review Template", "Review Presentation", and "Adapted Metrics" have received little attention from researchers so far.[7]

Review Template: One central challenge addressed by studies is how review quality can be improved by providing reviewers with either general guidelines (e.g., Cao et al., 2011) or a review template (e.g., Yin et al., 2014). The key idea behind this approach is that once the drivers of the perceived helpfulness and general review quality are known, system designers can leverage these drivers by alluding to them in review templates or guidelines.

Review Presentation: Prior research has also pointed out that the order in which systems present reviews to consumers can have crucial implications. Authors of these studies argue that the order of the reviews presented has a moderating effect on the direct outcome of online reviews on sales, for instance (e.g., Hu et al., 2014).

Adapted Metrics: Most review systems provide aggregate information on all given reviews for a product in form of metrics such as volume and valence. Yet, many studies argue that the positive direct outcome effects of online reviews can be enhanced by providing additional metrics (e.g., Hu et al., 2017a).

New design features. Even though there is a considerable backlog in the literature of proposed but not yet analyzed design features, new and unknown design features moderating driver or outcome effects of online reviews also represent fertile ground for future research. For example, one recent study investigating the effect of displaying product-related questions and answers (Q&A) on the valence of the subsequent ratings and sales (Banerjee et al., 2017b) suggests that displaying Q&As reduces the fit uncertainty between the consumer and the product. Further investigating the role of Q&A on a review system could enhance our current understanding of

---

[7] In Table A6 in the appendix, we provide a list of proposed design features for all the design feature categories identified in our literature synthesis (cf. Table 1), including respective design features that have already been evaluated by researchers.

Table 3

List of proposed but not analyzed design features.

Design feature category	Proposed design feature	Proposed in articles
Review template	Design review templates so that reviewers provide more information on product quality/experiences	Kwark et al. (2018)Ludwig et al. (2013)Raguseo et al. (2017)
	Provide guidelines for writing more helpful reviews	Cao et al. (2011)Hu et al. (2014)Yin et al. (2014)Racherla and Friske (2012)Kuan et al. (2015)
	Encourage users to report past purchasing behavior and usage activity	Cheung et al. (2012)Cheung et al. (2014)
Review presentation	Change phrasing of question about helpfulness	Yin et al. (2016)
	Promote most helpful reviews	Hu et al. (2014)Ngo-Ye and Sinha (2014)
	Show information on the reviewer's background (e.g., purchasing history)	Liu and Karahanna (2017)
	Display reviews in decreasing order of reviewer trustworthiness/reputation	Banerjee et al. (2017a)Zhou and Guo (2017)Cheung et al. (2012)
	Present average rating next to each individual rating	Qiu et al. (2012)
	Enable users to report uninformative and non-persuasive reviews	Zhang et al. (2014)
	Differentiated helpfulness score	Yin et al. (2017)Scholz and Dorner (2013)
Adapted metrics	Display variance of online ratings	He and Bond (2015)Wang et al. (2015)
	Display distribution of online ratings	Hu et al. (2017a)
	Display number of transactions	Dellarocas and Wood (2008)
	Display reputation scores per category	Kokkodis and Ipeirotis (2016)
	Make volume less prominent	Dellarocas et al. (2010)
	Let old reviews expire	Jin et al. (2014)
	Let negative reviews expire	Moreno and Terwiesch (2014)
	Do not delete/censor negative reviews	Lee et al. (2017)Pan and Zhang (2011)
	Place higher weighting on recent transactions	Ghose (2009)
	Predict and display long-term average rating	Khopkar and Nikolaev (2017)
	Weight reviews based on cultural background	Hong et al. (2016)

product fit uncertainty (e.g., Hong and Pavlou, 2014). Given the number of new design features that online review platforms are experimenting with, we recognize some potential for future research to investigate these design features. For example, Yelp recently started to introduce summary statistics of the geographic locations where a reviewer has reviewed businesses. This design feature addresses the relationship between reviewer origin – i.e., where a reviewer currently resides – and online reviewing behavior (Yelp, 2017b).

Environments

Based on the extended coding (n = 58 articles), we found that the majority of studies analyze design features in B2C (79%) and C2C (19%) environments. However, online review platforms such as TrustRadius or G2Crowd illustrate that online reviews are also used in B2B environments. Moreover, the majority of existing studies analyze one-sided review systems (95%) where only the buyer reviews the product or service. However, in recent years, platform-based business models have been established by firms such as Airbnb or Uber that orchestrate two-sided markets (e.g., Parker and van Alstyne, 2005). In such environments, two-sided review systems where buyers and sellers mutually review each other are implemented. Finally, the majority of studies (97%) analyze reviewing behavior in a pure online environment. However, consumers often search for product information online and then make the purchase offline, the phenomenon of research shopping (Verhoef et al., 2007). Such an online-offline-interplay is especially important for stationary local businesses. Hence, we state our second research gap.

Research Gap 2: The design of review systems for different environments is under-investigated.

B2B environments. Review systems for B2B environments pose interesting challenges for research on the design of such systems and general research on online reviews. Studies have quantitatively and qualitatively analyzed the drivers of online reviews primarily in a retail consumer context (e.g., Hennig-Thurau et al., 2004, Li and Hitt, 2008). In a business context, however, the reviewer is not a single consumer but represents a business, which is likely to have a bearing on the drivers of online reviews. Publishing a review in a B2B setting is associated with complex strategic decisions. Businesses that are keen to preserve a competitive advantage over competitors might have no interest in improving the reputation of their own suppliers, for example. Therefore, depending on the strategic importance of the business or product to be reviewed, businesses might apply different strategies to online reviews. Furthermore, the writing style in textual reviews could be less driven by emotions (Elfenbein, 2007), as reviewers are more likely to argue on a professional level with codified language associated with a particular business domain.

The economic impact of B2B reviews is equally unclear. While the positive effect of online reviews on sales for B2C relationships has been thoroughly investigated (e.g., Babic Rosario et al., 2016), the question remains whether and how an improvement in online reputation influences firm performance for B2B businesses.

Two-sided review systems. Two-sided reviewing can introduce a positive reciprocity bias to online reviews due to a fear of retaliation by one of the two parties or due to abstention bias (e.g., Bolton et al., 2013). Studies conclude that anonymous one-sided or double-

blind reviewing can remedy the reciprocation bias (Bolton et al., 2013). For many sharing economy platforms, this is not a viable option because both parties want to acquire information on each other. That is, guests on Airbnb would like to know who the prior guests were and what they have said, and hosts would like to know what prior hosts have said about their potential guests.

Another study finds that designers of review systems need to find ways of alleviating racial discrimination in sharing markets that disfavor some ethnic groups, e.g. to be hosted as guests by Airbnb hosts (Edelman et al., 2017). To alleviate this bias, the authors suggest anonymous booking requests. For the review system designer this means deciding whether to trade off the erosion of information asymmetry for privacy-preserving measures, i.e., anonymous booking requests. However, as of now this trade-off in the design of review systems is only insufficiently explored.

Finally, a frequently used way to monetize on transactions between two market sides on a platform is to ask for fees for each transaction from both sides, which are often asymmetric. This pricing mechanism challenges the applicability of the existing results from one-sided review systems to two-sided ones. Review system designers of two-sided platforms have to implement design features that support the supply side, such as reputation profile updating based on ratings in a certain time window (Aperjis and Johari, 2010) or withholding the display of negative reviews (Chen and Xie, 2008). At the same time, design features should also help the demand side – such as displaying either attribute-oriented or usage-oriented reviews (Li et al., 2017), or highlighting negative reviews (Ein-Gar et al., 2012). Both of these – sometimes conflicting – objectives have to be considered simultaneously in two-sided review system design but remain unexplored so far.

Online-offline-interplay. Only very few studies have examined the impact of the local offline environment on online review behavior. One such study that has done so (Huang et al., 2016) suggests that due to the systematic impact of geographic heterogeneity on the valence of online ratings, review systems should facilitate the identification of a user's home area in order to enable future consumers to account for this systematic rating difference. Yet, this design feature has not yet been analyzed, even though Yelp, for instance, has recently introduced measures that allow to track the statistics of where reviewers have posted their reviews (Yelp, 2017b).

In another example, Amazon has recently started to launch local bricks-and-mortar stores that display the online ratings on their shelves (Amazon, 2017). This practice poses a whole new challenge to our existing knowledge on the effects of online reviews, as it is not known whether local offline search behavior in a physical store significantly differs from how consumers search online.

Devices

Based on our extended coding, we found that none of the existing studies analyze design features of review systems for mobile consumers.[8] However, online shopping behavior has drastically changed in recent years towards a more mobile and tablet-based economy (e.g., Ghose, 2017). Statistics issued by review platforms also point to the fact that more and more reviews are given via mobile devices, while those given via stationary devices are stagnating (Yelp, 2017a). Similarly, speech-controlled devices like Amazon's Echo are rapidly gaining in popularity, and these may fundamentally alter the review production and consumption behavior of consumers. Similarly, augmented reality devices should challenge the conventional knowledge on review production and consumption behavior and the applications running on these devices (Fang et al., 2014). Accordingly, we state our third research gap.

Research Gap 3: The design of review systems for different devices is under-investigated.

Mobile devices. Since the advent of mobile shopping, consumers can access – as well as write and upload – online reviews anytime and anywhere. Using a mobile device has a number of implications for online review system design, as it changes how reviews are generated and consumed. For example, reading reviews on a smaller screen when out shopping could result in consumers reading fewer reviews. And whilst a mobile device makes it easier to review a service soon after experiencing it, it has been found that instant reviewing negatively affects the valence of online reviews (Huang et al., 2016) but positively affects their perceived helpfulness (Chen and Lurie, 2013). Writing reviews on a smaller screen may shorten the textual part of the review but enhance the value – and perceived helpfulness – of the review through the uploading of photos or videos illustrating the product or business. Writing on a smaller screen also increases the usage of informal speech in review texts (Ransbotham et al., 2018) – as has been the case with mobile applications such as WhatsApp, Twitter, or Facebook messenger – which in turn may have implications that have not yet been researched so far.

Speech-controlled devices. Virtually all the studies we reviewed rest upon the assumption that reviews are generated manually (i.e. being typed into a device) and visually searched for and read by potential consumers. This assumption has to be challenged as new technologies, such as Amazon's speech-controlled device Echo, will sooner or later change the way people shop for items online and how reviews, in turn, are written and consumed. Equally, not all consumers may adopt speech enabled technology, thus further increasing the challenge of meeting the needs of different consumer sections.

A number of findings that have investigated the drivers and outcomes of reviews may have to be revised and new research conducted, if people change the way they search for and compare different product offers. For example, some studies have demonstrated that the (visual) display format and the aggregation level of reviews can act as a driver of reviews (e.g., Chen et al., 2018) which in turn can affect economic outcomes (e.g., Li, 2017), but what would be the impact if reviews are consumed via speech controlled devices, without the visual display on which online review systems currently rely?

---

[8] We note that Karimi and Wang (2017) analyze online reviews for mobile app games, which are probably given via mobile devices. Yet, the authors do not focus on the particularities introduced via this review channel.

Likewise, generating reviews with speech-enabled technology would increase the spoken (and informal) language being used in reviews. While the cost (time and effort) of posting a review might decrease, the cost of attaching additional information, such as photos, might increase. Thus, not only the volume but also the shape and format of a review might change due to new technology, which will require adjustments to be made to online review system design.

Research agenda – How do we get there?

In the following, we develop a research agenda to help scholars address the research gaps presented in Section 'Research gaps'.

Research on design features

In Section 'Design features', we have presented studies that have proposed a range of design features, grouped into categories (see Table 3). To address our first research gap, we suggest the following overarching research thrust:
Research Thrust 1: *What are the moderating effects of already proposed (but not yet analyzed) and new design features?*

Theoretical viability
Any changes to the design of an online review system should be aligned to the system owner's (i.e., third-party platform or seller) digital strategy (Yeow et al., 2018). Every incremental design decision moderates a direct outcome or a direct driver effect. Behind each of these effects, there are underlying theoretical mechanisms that determine the impact of the design change. Exploring these design changes can help advance these theories and deepen the understanding of how consumers interact with online review systems. Analyzing new design features such as review templates, for example, can give further insights into the writing process for tasks that involve creative writing as well as recalling an event, in this case, the consumption experience. Research focusing on design features that moderate the direct outcome effect bear the potential to advance the theoretical body of knowledge on the understanding of human information processing. A better understanding of how humans process the display of online rating information can, in turn, inform review system designers as to which design features to implement to support the strategic focus of the e-commerce platform. We propose Table 3 to be used as a starting point for this research area since it contains relevant research articles that suggest a multitude of design features to form the foundation of a new design theory (e.g., Gregor and Jones, 2007).

Methodological viability
Future research could use the list of proposed design features and then track down the changelog of review systems to identify policy changes enabling the evaluation of a proposed design feature. Studies using observational panel data seem primed for this task. In order to visually investigate design changes on the website, researchers can use tools such as the waybackmachine.org to compare how the website looked before and after the design change. Classical difference-in-difference analysis or regression discontinuity approaches (Angrist and Pischke, 2009) are obvious methodological choices in this case. For future research to investigate the effectiveness of new design features, analytical models and (field) experimental studies appear as the first choice.

Research on review system design for different environments

Besides pursuing the above research thrust in conventional B2C environments (e.g., with platforms like Amazon or Yelp), research can also study the moderating effects of design features in different environments. Thus, we formulate our second research thrust:
Research Thrust 2: *How do different environments change the moderating effects of design features and how do review systems need to be designed for different environments?*

Theoretical viability
The second research thrust can contribute to a design theory of review systems by highlighting that online reviewing behavior is (to some extent) dependent on the environment. Review system designers need to account for whether environmental dynamics (e.g., traveling, augmented reality games, etc.) might affect the online rating behavior of the product or service domain for which they are designing a review system. For example, accounting for information on the offline environment of reviewers might be more important for local services, such as restaurants, than for digital products such as computer software. For theoretical guidance on the influence of environmental factors, scholars can borrow theories (e.g., Truex et al., 2006) from cogent domains.
For B2C markets, online reviews help to build trust among market participants (e.g., Ba and Pavlou, 2002). In the B2B context which strongly differs from the B2C context in several aspects (e.g., McKnight et al., 2017), organizational trust has become an established subject of research (e.g., Mayer et al., 1995) and a relevant component of strategic decision-making (e.g., McKnight et al., 2017). Building on first results that present evidence for the trust-building capability of online reviews in the B2B context (Pavlou, 2002), and established models of organizational trust (e.g., McKnight et al., 2002), research on design features for B2B online review platforms can broaden the theoretical understanding of organizational trust. Moreover, task delegation issues may arise between principals and agents and review giving and reading could be delegated. Researchers could borrow theories from behavioral economics – such as the principal-agent theory (e.g., Holmstrom and Milgrom, 1991) – to theorize on design features for B2B contexts.
Similarly, future research on two-sided review environments could reveal extensions to theories related to social interaction (e.g., Proserpio et al., 2016). Primary research in the context of offline-online dynamics in online review systems has already provided an extension to the construal-level theory, for instance (Huang et al., 2016).

Methodological viability

For the analysis of B2B environments, some empirical evidence can be gleaned from observational data from platforms such as TrustRadius or G2Crowd. Additionally, while it might be difficult to mimic a B2B scenario in a classical lab experiment, scenario experiments might be conducted to investigate participants' online rating behavior where they are reading or writing a review on behalf of their company. Similarly, observational data and scenario experiments could be used to investigate the effects of online-offline dynamics. In many instances, review platforms such as TripAdvisor provide consumers with information on the origin of reviewers, and researchers can make use of this. Scenario experiments can also be used to induce participants to take the perspective of travelers or locals. Finally, classical lab experiments or policy changes from two- to one-sided review policies (or vice versa) seem to be the ideal research environment in which to investigate two-sided reviewing.

Research on review system design for different devices

In light of the ongoing increase in the usage of mobile, speech-controlled, and augmented reality devices scholars should study the design of review systems for different devices. Consequently, we suggest the following research thrust:

Research Thrust 3: *How do different devices change the moderating effects of design features and how do review systems need to be designed for different devices?*

Theoretical viability

The digitization of the world through new devices represents an important opportunity for the information systems community (Sørensen and Landau, 2015). Scholars and practitioners need to understand which design features can foster the adoption of novel internet devices to be used for reading and writing (or, listening to, and dictating) reviews. To this end, researchers can build on and contribute to theory of IT adoption (e.g., Davis, 1989) or even novel research streams like neuro IS (e.g., Dimoka et al., 2011) – in particular for augmented reality. Besides focusing on adoption, one can also focus on how new devices alter the way in which consumers interact with review systems. For instance, mobile devices give more freedom to users (Jarvenpaa et al., 2003) who have an increased opportunity to use the Internet anywhere (Ghose and Han, 2011). This technological innovation allows for changes in consumer behavior that can be studied to extend theoretical knowledge.

Methodological viability

Future work could observe and evaluate existing design differences between stationary and mobile versions of review systems to provide an overview of possible design features from a practitioner's perspective. Furthermore, studies could analyze existing research findings on the moderating effect of a single design feature with respect to differences between the mobile and the stationary use of the system. To evaluate any design feature with regards to its usefulness on mobile devices, scholars need to know which type (mobile or stationary) a reviewer has used to write their review. Therefore, we propose to (1) form partnerships with businesses hosting a review system, (2) examine review systems that reveal the reviewer's device (e.g., TripAdvisor), and (3) run lab and field experiments in which participants use different devices. While reviewing and reading reviews with mobile devices becomes increasingly common, using review systems with speech-controlled devices is, to the best of our knowledge, still non-existent. Anecdotal evidence suggests that as of now, Amazon's Echo, for instance, is rather awkward for shopping and is mainly used for repeat purchases of consumer items (Gikas, 2016). Therefore, future research first needs to establish theoretical foundations and an appropriate design theory enabling users with speech-controlled devices to adequately access review systems. The same holds for further potential new technologies such as augmented or virtual reality devices.

Discussion and conclusion

From the perspective of the resource-based view of the firm, review systems populated with online reviews constitute a specialized asset. Still, there are at least two fundamental differences.

First, specialized assets are usually created and used within the firm. Review systems, however, leverage a power shift from within the firm to outside the firm. E-commerce platforms running these systems source, accumulate, and aggregate consumption experience from people outside the firm (e.g., Raguseo et al., 2017) as their main providers of strategically important knowledge.

Second, as discussed in this article, online reviews may be collected, processed, aggregated, and presented in quite different ways and many design features have been identified in the literature that influence the drivers and economic outcomes of online reviews. It is the unique combination of design features of a specific review system that may increase its potential to become a specialized asset for the firm hosting it. But, unlike the customization of ERP systems (Beard and Sumner, 2004), implemented design features of review systems can be easily observed from outside the firm and may be imitated by rivals. Hence, competitive advantage will not directly stem from single design features but from the knowledge about their combined effects implemented on specific devices in specific environments, which are not easily observable by outsiders.

Consequently, future research on the design of review systems should always address a combination of the three research thrusts proposed in this literature review article: Research Thrust 1 accounts for potential design features that may be combined. Research Thrust 2 accounts for the growing diversity of environments in which review systems are applied. Research Thrust 3 is devoted to the fact that online reviews are not only created and consumed via stationary personal computers but through a multiplicity of (mobile) devices, which require different design choices. Consequently, answers to the following synthesized question should inform further research:

Which design feature combinations should be implemented and employed on different devices given the chosen environment of an e-commerce platform?

These answers may be used to align strategic goals with the design of the review system. An e-commerce platform may implement different strategies such as focusing on increasing the number of reviews (volume) or the quality of online reviews (e.g., better consumer learning, perceived helpfulness,). Accordingly, e-commerce platforms have to align their review systems to the platform strategy and choose review system design features accordingly. For instance, an e-commerce platform can implement a combination of design features to attract a high number of reviews. If a review system fails to accumulate a critical mass of reviews, the risk of the system to fail is substantial. Other e-commerce platforms may put their strategic focus on the consumers' quality perception of the review system. Accordingly, such platforms should combine design features incentivizing reviewers to post high quality online reviews with features that help users identify these high quality reviews.

Moreover, the strategic goals of e-commerce platforms may change over time. For instance, an e-commerce startup primarily needs a critical mass of online reviews to attract users. Consequently, such a platform may follow the strategy to increase the number of reviews first. As soon as a critical mass is reached, it may be beneficial to adapt the strategy towards increasing the quality of reviews, and align the review system design accordingly. Consequently, aligning the review system with different strategies by adding and removing design features is a dynamic and ongoing process rather than a static one-shot decision.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jsis.2019.01.004.

References

Amazon, 2017. Amazon Books. URL: < https://www.amazon.com/b?node=13270229011 > (visited on 07/13/2018).
Angrist, J., Pischke, J.-S., 2009. Mostly Harmless Econometrics: An Empiricists Guide. Princeton University Press, Princeton.
Aperjis, C., Johari, R., 2010. Optimal windows for aggregating ratings in electronic marketplaces. Manage. Sci. 56 (5), 864–880.
Askalidis, G., Kim, S.J., Malthouse, E.C., 2017. Understanding and overcoming biases in online review systems. Decis. Support Syst. 97, 23–30.
Avery, C., Resnick, P., Zeckhauser, R., 1999. The market for evaluations. Am. Econ. Rev. 89 (3), 564–584.
Ba, S., Pavlou, P.A., 2002. Evidence of the effect of trust building technology in electronic markets. Price premiums and buyer behavior. MIS Quart. 26 (3), 243.
Babic Rosario, A., Sotgiu, F., de Valck, K., Bijmolt, T.H.A., 2016. The effect of electronic word of mouth on sales. A meta-analytic review of platform, product, and metric factors. J. Mark. Res. 53 (3), 297–318.
Banerjee, S., Bhattacharyya, S., Bose, I., 2017a. Whose online reviews to trust? Understanding reviewer trustworthiness and its impact on business. Decis. Support Syst. 96, 17–26.
Banerjee, S., Dellarocas, C., Zervas, G., 2017b. Interacting User Generated Content Technologies: How Q&As Affect Ratings & Reviews. ACM Economics & Computation.
Barney, J., 1991. Firm resources and sustained competitive advantage. J. Manage. 1 (17), 99–120.
Beard, J.W., Sumner, M., 2004. Seeking strategic advantage in the post-net era: viewing ERP systems from the resource-based perspective. J. Strategic Inform. Syst. 13 (2), 129–150.
Benlian, A., Titah, R., Hess, T., 2012. Differential effects of provider recommendations and consumer reviews in E-commerce transactions. An experimental study. J. Manage. Inform. Syst. 29 (1), 237–272.
Bharadwaj, A.S., 2000. A resource-based perspective on information technology capability and firm performance: an empirical investigation. MIS Quart. 24 (1), 169–196.
Bolton, G., Greiner, B., Ockenfels, A., 2013. Engineering trust: reciprocity in the production of reputation information. Manage. Sci. 59 (2), 265–285.
Bolton, G., Greiner, B., Ockenfels, A., 2018. Dispute resolution or escalation? The strategic gaming of feedback withdrawal options in online markets. Manage. Sci. 64 (9), 3971–4470.
Burtch, G., Hong, Y., 2014. What happens when word of mouth goes mobile?. Proceedings of the Thirty Fifth International Conference on Information Systems, Auckland, New Zealand.
Camilleri, A.R., 2017. The presentation format of review score information influences consumer preferences through the attribution of outlier reviews. J. Interactive Market. 39, 1–14.
Cao, Q., Duan, W., Gan, Q., 2011. Exploring determinants of voting for the "helpfulness" of online user reviews. A text mining approach. Decis. Support Syst. 50 (2), 511–521.
Chang, H.H., Wu, L.H., 2014. An examination of negative e-WOM adoption. Brand commitment as a moderator. Decis. Support Syst. 59, 206–218.
Chang, M.K., Cheung, W., Tang, M., 2013. Building trust online: interactions among trust building mechanisms. Inform. Manage. 50 (7), 439–445.
Chen, L., Jiang, T., Li, W., Geng, S., Hussain, S., 2017. Who should pay for online reviews? Design of an online user feedback mechanism. Electron. Commer. Res. Appl. 23, 38–44.
Chen, P.-Y., Hong, Y., Liu, Y., 2018. The value of multidimensional rating systems: evidence from a natural experiment and randomized experiments. Manage. Sci. 64 (10), 4629–4647.
Chen, Y., Harper, M.F., Konstan, J., Li, S.X., 2010. Social comparison and contribution to online communities: a field experiment on movielens. Am. Econ. Rev. 100 (4), 1358–1398.
Chen, Y., Xie, J., 2008. Online consumer review. word-of-mouth as a new element of marketing communication mix. Manage. Sci. 54 (3), 477–491.
Chen, Z., Lurie, N.H., 2013. Temporal contiguity and negativity bias in the impact of online word of mouth. J. Mark. Res. 50 (4), 463–476.
Cheung, C., Sia, C.-L., Kuan, K., 2012. Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. J. Assoc. Inform. Syst. 13 (8), 618–635.
Cheung, C.M.K., Thadani, D.R., 2012. The impact of electronic word-of-mouth communication. A literature analysis and integrative model. Decis. Support Syst. 54 (1),

461–470.

Cheung, C.M.K., Xiao, B.S., Liu, I.L.B., 2014. Do actions speak louder than voices? The signaling role of social information cues in influencing consumer purchase decisions. Decis. Support Syst. 65, 50–58.

Cheung, K.-W., Kwok, J.T., Law, M.H., Tsui, K.-C., 2003. Mining customer product ratings for personalized marketing. Decis. Support Syst. 35 (2), 231–243.

Chevalier, J.A., Mayzlin, D., 2006. The effect of word of mouth on sales. Online book reviews. J. Mark. Res. 43 (3), 345–354.

Chung, C.-Y., Hsu, P.-Y., Huang, S.-H., 2013. A novel approach to filter out malicious rating profiles from recommender systems. Decis. Support Syst. 55 (1), 314–325.

Davis, F.D., 1989. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quart. 13 (3), 319–340.

Dellarocas, C., 2003. The digitization of word of mouth. Promise and challenges of online feedback mechanisms. Manage. Sci. 49 (10), 1407–1424.

Dellarocas, C., 2005. Reputation mechanism design in online trading environments with pure moral hazard. Inform. Syst. Res. 16 (2), 209–230.

Dellarocas, C., 2006. How often should reputation mechanisms update a trader's reputation profile? Inform. Syst. Res. 17 (3), 271–285.

Dellarocas, C., Gao, G., Narayan, R., 2010. Are consumers more likely to contribute online reviews for hit or niche products? J. Manage. Inform. Syst. 27 (2), 127–158.

Dellarocas, C., Wood, C.A., 2008. The sound of silence in online feedback: estimating trading risks in the presence of reporting bias. Manage. Sci. 54 (3), 460–476.

Dimoka, A., Pavlou, P.A., Davis, F.D., 2011. NeuroIS: the potential of cognitive neuroscience for information systems research. Inform. Syst. Res. 22 (4), 687–702.

Duan, W., Gu, B., Whinston, A.B., 2008. Do online reviews matter? — an empirical investigation of panel data. Decis. Support Syst. 45 (4), 1007–1016.

Edelman, B., Luca, M., Svirsky, D., 2017. Racial discrimination in the sharing economy. Evidence from a field experiment. Am. Econ. J.: Appl. Econ. 9 (2), 1–22.

Ein-Gar, D., Shiv, B., Tormala, Z.L., 2012. When blemishing leads to blossoming. The positive effect of negative information. J. Consum. Res. 38 (5), 846–859.

Elfenbein, H.A., 2007. 7 Emotion in organizations. Acad. Manage. Ann. 1 (1), 315–386.

Fang, H., Zhang, J., Şensoy, M., Magnenat-Thalmann, N., 2014. Reputation mechanism for E-commerce in virtual reality environments. Electron. Commer. Res. Appl. 13 (6), 409–422.

Floyd, K., Freling, R., Alhoqail, S., Cho, H.Y., Freling, T., 2014. How online product reviews affect retail sales. A meta-analysis. J. Retail. 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. Inform. Syst. Res. 19 (3), 291–313.

Gable, G., 2010. Strategic information systems research. An archival analysis. J. Strat. Inform. Syst. 19 (1), 3–16.

Ghose, A., 2009. Internet exchanges for used goods. An empirical analysis of trade patterns and adverse selection. MIS Quart. 33 (2), 263–291.

Ghose, A., 2017. Tap: Unlocking the Mobile Economy. MIT Press.

Ghose, A., Han, S.P., 2011. An empirical analysis of user content generation and usage behavior on the mobile Internet. Manage. Sci. 57 (9), 1671–1691.

Ghose, A., Ipeirotis, P.G., Li, B., 2012. Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. Market. Sci. 31 (3), 493–520.

Gikas, M., 2016. What Amazon Echo and Alexa do Best. URL: < https://www.consumerreports.org/wireless-speakers/what-amazon-echo-and-alexa-do-best/ > (visited on 07/13/2018).

Goes, P.B., Lin, M., Au Yeung, C.-M., 2014. "Popularity Effect" in user-generated content. Evidence from online product reviews. Inform. Syst. Res. 25 (2), 222–238.

Gregor, S., Jones, D., 2007. The anatomy of a design theory. J. Assoc. Inform. Syst. 8 (5), 312–335.

Gu, B., Ye, Q., 2014. First step in social media. Measuring the influence of online management responses on customer satisfaction. Prod. Oper. Manage. 23 (4), 570–582.

He, S.X., Bond, S.D., 2015. Why is the crowd divided? Attribution for dispersion in online word of mouth. J. Consum. Res. 41 (6), 1509–1527.

Hennig-Thurau, T., Gwinner, K.P., Walsh, G., Gremler, D.D., 2004. Electronic word-of-mouth via consumer-opinion platforms. What motivates consumers to articulate themselves on the internet? J. Interactive Market. 18 (1), 38–52.

Holmstrom, B., Milgrom, P., 1991. Multitask principal-agent analyses: incentive contracts, asset ownership, and job design. JL Econ. Org. 7, 24.

Hong, H., Di, Xu., Wang, G.A., Fan, W., 2017. Understanding the determinants of online review helpfulness. A meta-analytic investigation. Decis. Support Syst. 102, 1–11.

Hong, Y., Huang, N., Burtch, G., Li, C., 2016. Culture, conformity, and emotional suppression in online reviews. J. Assoc. Inform. Syst. 17 (11), 737–758.

Hong, Y., Pavlou, P.A., 2014. Product fit uncertainty in online markets. Nature, effects, and antecedents. Inform. Syst. Res. 25 (2), 328–344.

Hostler, R.E., Yoon, V.Y., Guo, Z., Guimaraes, T., Forgionne, G., 2011. Assessing the impact of recommender agents on on-line consumer unplanned purchase behavior. Inform. Manage. 48 (8), 336–343.

Hu, N., Bose, I., Koh, N.S., Liu, L., 2012. Manipulation of online reviews. An analysis of ratings, readability, and sentiments. Decis. Support Syst. 52 (3), 674–684.

Hu, N., Koh, N.S., Reddy, S.K., 2014. Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales. Decis. Support Syst. 57, 42–53.

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

Hu, Y.-H., Chen, K., Lee, P.-J., 2017b. The effect of user-controllable filters on the prediction of online hotel reviews. Inform. Manage. 54 (6), 728–744.

Huang, L., Tan, C.-H., Ke, W., Wei, K.-K., 2014. Do we order product review information display? How? Inform. Manage. 51 (7), 883–894.

Huang, N., Burtch, G., Hong, Y., Polman, E., 2016. Effects of multiple psychological distances on construal and consumer evaluation. A field study of online reviews. J. Consum. Psychol. 26 (4), 474–482.

Jabr, W., Zheng, E., 2014. Know yourself and know your enemy. An analysis of firm recommendations and consumer reviews in a competitive environment. MIS Quart. 38 (3), 635–654.

Jarvenpaa, S.L., Lang, K.R., Takeda, Y., Tuunainen, V.K., 2003. Mobile commerce at crossroads. Commun. ACM 46 (12), 41.

Jiang, Y., Guo, H., 2015. Design of consumer review systems and product pricing. Inform. Syst. Res. 26 (4), 714–730.

Jin, L., Hu, B., He, Y., 2014. The recent versus the out-dated. An experimental examination of the time-variant effects of online consumer reviews. J. Retail. 90 (4), 552–566.

Karimi, S., Wang, F., 2017. Online review helpfulness. Impact of reviewer profile image. Decis. Support Syst. 96, 39–48.

Khopkar, S.S., Nikolaev, A.G., 2017. Predicting long-term product ratings based on few early ratings and user base analysis. Electron. Commer. Res. Appl. 21, 38–49.

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. J. Interactive Market. 28 (3), 167–183.

Koh, N.S., Hu, N., Clemons, E.K., 2010. Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. Electron. Commer. Res. Appl. 9 (5), 374–385.

Kokkodis, M., Ipeirotis, P.G., 2016. Reputation transferability in online labor markets. Manage. Sci. 62 (6), 1687–1706.

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. Electron. Commer. Res. Appl. 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. J. Assoc. Inform. Syst. 16 (1), 48–71.

Kwark, Y., Chen, J., Raghunathan, S., 2018. User-generated content and competing firms' product design. Manage. Sci. 64 (10), 4608–4628.

Lee, J.H., Jung, S.H., Park, J., 2017. The role of entropy of review text sentiments on online WOM and movie box office sales. Electron. Commer. Res. Appl. 22, 42–52.

Li and Hitt, 2010. Price effects in online product review: an analytical model and empirical analysis.

Li, L., Xiao, E., 2014. Money talks. Rebate mechanisms in reputation system design. Manage. Sci. 60 (8), 2054–2072.

Li, M., Tan, C.-H., Wei, K.-K., Wang, K., 2017. Sequentiality of product review information provision. An information foraging perspective. Manage. Inform. Syst. Quart. 41 (3), 867–892.

Li, X., 2017. Revealing or non-revealing. The impact of review disclosure policy on firm profitability. MIS Quart. 41 (4), 1335–1345.

Li, X., Hitt, L.M., 2008. Self-selection and information role of online product reviews. Inform. Syst. Res. 19 (4), 456–474.

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

Ludwig, S., de Ruyter, K., Friedman, M., Brüggen, E.C., Wetzels, M., Pfann, G., 2013. More than words. The influence of affective content and linguistic style matches in online reviews on conversion rates. J. Market. 77 (1), 87–103.

de Matos, C.A., Rossi, C.A.V., 2008. Word-of-mouth communications in marketing. A meta-analytic review of the antecedents and moderators. J. Acad. Mark. Sci. 36 (4), 578–596.

Mayer, R.C., Davis, J.H., Schoorman, F.D., 1995. An integrative model of organizational trust. Acad. Manag. Rev. 20 (3), 709–734.

Mayzlin, D., Dover, Y., Chevalier, J., 2014. Promotional reviews. An empirical investigation of online review manipulation. Am. Econ. Rev. 104 (8), 2421–2455.

McKnight, D.H., Lankton, N.K., Nicolaou, A., Price, J., 2017. Distinguishing the effects of B2B information quality, system quality, and service outcome quality on trust and distrust. J. Strateg. Inform. Syst. 26 (2), 118–141.

McKnight, D.H., Choudhury, V., Kacmar, C., 2002. The impact of initial consumer trust on intentions to transact with a web site. A trust building model. J. Strateg. Inform. Syst. 11 (3–4), 297–323.

Moreno, A., Terwiesch, C., 2014. Doing business with strangers. Reputation in online service marketplaces. Inform. Syst. Res. 25 (4), 865–886.

Muchnik, L., Aral, S., Taylor, S.J., 2013. Social influence bias: a randomized experiment. Science 341 (6146), 647–651.

Neumann, J., Gutt, D., 2017. A homeowner's guide to airbnb. Theory and empirical evidence for optimal pricing conditional on online ratings. In: Proceedings of the 25th European Conference on Information Systems (ECIS), pp. 997–1010.

Ngo-Ye, T.L., Sinha, A.P., 2014. The influence of reviewer engagement characteristics on online review helpfulness. A text regression model. Decis. Support Syst. 61, 47–58.

Ong, T., Mannino, M., Gregg, D., 2014. Linguistic characteristics of shill reviews. Electron. Commer. Res. Appl. 13 (2), 69–78.

Pan, Y., Zhang, J.Q., 2011. Born unequal. A study of the helpfulness of user-generated product reviews. J. Retail. 87 (4), 598–612.

Panniello, U., Gorgoglione, M., Tuzhilin, A., 2016. Research note—In CARSs we trust. How context-aware recommendations affect customers' trust and other business performance measures of recommender systems. Inform. Syst. Res. 27 (1), 182–196.

Paré, G., Trudel, M.-C., Jaana, M., Kitsiou, S., 2015. Synthesizing information systems knowledge. A typology of literature reviews. Inform. Manage. 52 (2), 183–199.

Parker, G.G., van Alstyne, M.W., 2005. Two-sided network effects. A theory of information product design. Manage. Sci. 51 (10), 1494–1504.

Pavlou, P.A., 2002. Institution-based trust in interorganizational exchange relationships. The role of online B2B marketplaces on trust formation. J. Strateg. Inform. Syst. 11 (3–4), 215–243.

Pavlou, P.A., Dimoka, A., 2006. The nature and role of feedback text comments in online marketplaces. Implications for trust building, price premiums, and seller differentiation. Inform. Syst. Res. 17 (4), 392–414.

Piramuthu, S., Kapoor, G., Zhou, W., Mauw, S., 2012. Input online review data and related bias in recommender systems. Decis. Support Syst. 53 (3), 418–424.

Proserpio, D., Xu, W., Zervas, G., 2016. You Get What You Give. Theory and Evidence of Reciprocity in the Sharing Economy. URL: https://pdfs.semanticscholar.org/6b7b/1238ed2a6915bee49dfbf9f9aa2454570f3e.pdf.

Proserpio, D., Zervas, G., 2015. Online Reputation Management: Estimating the Impact of Management Responses on Consumer Reviews.

Qiu, L., Pang, J., Lim, K.H., 2012. Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity. The moderating role of review valence. Decis. Support Syst. 54 (1), 631–643.

Racherla, P., Friske, W., 2012. Perceived 'Usefulness' of online consumer reviews. An exploratory investigation across three services categories. Electron. Commer. Res. Appl. 11 (6), 548–559.

Raguseo, E., Neirotti, P., Paolucci, E., 2017. How small hotels can drive value their way in infomediation. The case of 'Italian Hotels vs. OTAs and TripAdvisor'. Inform. Manage. 54 (6), 745–756.

Ransbotham, S., Lurie, N.H., Liu, H., 2018. Creation and consumption of mobile word-of-mouth: how are mobile reviews different? Marketing Sci (Articles in Advance).

Roberts, N., Galluch, P.S., Dinger, M., Grover, V., 2012. Absorptive capacity and information systems research: review, synthesis, and directions for future research. Manage. Inform. Syst. Quart. 36 (3).

Samak, A.C., 2013. An experimental study of reputation with heterogeneous goods. Decis. Support Syst. 54 (2), 1134–1149.

Schlosser, A.E., 2011. Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments. J. Consumer Psychol. 21 (3), 226–239.

Scholz, M., Dorner, V., 2013. The recipe for the perfect review? Bus. Inform. Syst. Eng. 5 (3), 141–151.

Shen, W., Hu, Y., Ulmer, J.R., 2015. Competing for attention: an empirical study of online reviewers' strategic behavior. MIS Quart. 39 (3), 683–696.

Sørensen, C., Landau, J.S., 2015. Academic agility in digital innovation research: the case of mobile ICT publications within information systems 2000–2014. J. Strateg. Inf. Syst. 24 (3), 158–170.

Spool, J., 2009. The Magic behind Amazon's 2.7 Billion Dollar Question. URL: https://articles.uie.com/magicbehindamazon/ (visited on 07/13/2018).

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

Sun, Y., Dong, X., McIntyre, S., 2017. Motivation of user-generated content. social connectedness moderates the effects of monetary rewards. Marketing Sci. 36 (3), 329–337.

Truex, D., Holmström, J., Keil, M., 2006. Theorizing in information systems research: a reflexive analysis of the adaptation of theory in information systems research. J. Assoc. Inform. Syst. 7 (1), 797–821.

van Noort, G., Willemsen, L.M., 2012. Online damage control: the effects of proactive versus reactive webcare interventions in consumer-generated and brand-generated platforms. J. Interactive Market. 26 (3), 131–140.

Verhoef, P.C., Neslin, S.A., Vroomen, B., 2007. Multichannel customer management: understanding the research-shopper phenomenon. Int. J. Res. Mark. 24 (2), 129–148.

Wang, F., Liu, X., Fang, E., 2015. User reviews variance, critic reviews variance, and product sales. An exploration of customer breadth and depth effects. J. Retail. 91 (3), 372–389.

Wang, Z., Li, H., Ye, Q., Law, R., 2016. Saliency effects of online reviews embedded in the description on sales. Moderating role of reputation. Decis. Support Syst. 87, 50–58.

Wu, C., Che, H., Chan, T.Y., Lu, X., 2015. The economic value of online reviews. Marketing Sci. 34 (5), 739–754.

Xu, P., Chen, L., Santhanam, R., 2015. Will video be the next generation of e-commerce product reviews? Presentation format and the role of product type. Decis. Support Syst. 73, 85–96.

Yang, J., Hu, X., Zhang, H., 2007. Effects of a reputation feedback system on an online consumer-to-consumer auction market. Decis. Support Syst. 44 (1), 93–105.

Ye, S., Gao, G., Viswanathan, S., 2014. Strategic behavior in online reputation systems. Evidence from eBay. MIS Quart. 38 (4), 1033–1056.

Yelp, 2017a. An Introduction to Yelp Metrics as of December 31, 2017. URL: https://www.yelp.com/factsheet (visited on 07/13/2018).

Yelp, 2017b. The Most Locals-Only Restaurants in San Francisco and San Diego. URL: https://www.yelpblog.com/2017/09/locals-restaurants-san-francisco-san-diego.

Yeow, A., Soh, C., Hansen, R., 2018. Aligning with new digital strategy: a dynamic capabilities approach. J. Strategic Inform. Syst. 27 (1), 43–58.

Yin, D., Bond, S., Zhang, H., 2014. Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. MIS Quart. 38 (2), 539–560.

Yin, D., Bond, S.D., Zhang, H., 2017. Keep your cool or let it out. Nonlinear effects of expressed arousal on perceptions of consumer reviews. J. Mark. Res. 54 (3), 447–463.

Yin, D., Mitra, S., Zhang, H., 2016. Research note—when do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth. Inform. Syst. Res. 27 (1), 131–144.

Ying, Y., Feinberg, F., Wedel, M., 2006. Leveraging missing ratings to improve online recommendation systems. J. Mark. Res. 43 (3), 355–365.

You, Y., Vadakkepatt, G.G., Joshi, A.M., 2015. A meta-analysis of electronic word-of-mouth elasticity. J. Market. 79 (2), 19–39.

Zacharia, G., Moukas, A., Maes, P., 2000. Collaborative reputation mechanisms for electronic marketplaces. Decis. Support Syst. 29 (4), 371–388.

Zhang, Y., Feick, L., Mittal, V., 2014. How males and females differ in their likelihood of transmitting negative word of mouth. J. Consumer Res. 40 (6), 1097–1108.

Zhou, S., Guo, B., 2017. The order effect on online review helpfulness. A social influence perspective. Decis. Support Syst. 93, 77–87.