





Decide Now or Later: Making Sense of Incoherence Across Online Reviews

Dezhi Yin,^{a,*} Triparna de Vreede,^a Logan M. Steele,^a Gert-Jan de Vreede^a

^aMuma College of Business, University of South Florida, Tampa, Florida 33620

*Corresponding author

Contact: dezhiyin@usf.edu,  <https://orcid.org/0000-0003-1107-3232> (DY); tdevreede@usf.edu,  <https://orcid.org/0000-0002-8037-8315> (TdV); lmsteele@usf.edu,  <https://orcid.org/0000-0002-1377-3147> (LMS); gdevreede@usf.edu,  <https://orcid.org/0000-0001-6909-9836> (G-JdV)

Received: September 3, 2021

Revised: March 15, 2022

Accepted: May 25, 2022

Published Online in Articles in Advance:
July 5, 2022

<https://doi.org/10.1287/isre.2022.1150>

Copyright: © 2022 INFORMS

Abstract. Mixed or inconsistent opinions are commonplace in online reviews. Prior research shows that review inconsistency has different effects: its product-level manifestation in the form of inconsistent product ratings is associated with poorer sales, but its review-level manifestation in the form of two-sided arguments is associated with greater helpfulness and credibility evaluations of the review. In practice, consumers rarely consult all reviews or just a single review before they make purchase decisions. Instead, they often read a set of featured reviews (i.e., a review set) and some additional reviews if needed. Focusing on inconsistency in a review set, we introduce a new type of inconsistency across reviews that does not exist within a single review or at the level of product ratings: cross-review incoherence, which refers to disagreement among reviewers about specific product features. Based on cognitive dissonance theory, we explore how and when cross-review incoherence influences helpfulness and credibility judgments of the review set, revealing that consumers' reactions to such incoherence might operate differently beyond an individual review. In addition, we examine how and why consumers' judgments of a review set influence their purchase deferral—that is, the likelihood of making a buy-or-not-buy decision immediately after consulting the top reviews or deferring it until after obtaining more information. Two laboratory experiments demonstrate that cross-review incoherence increases purchase deferral via more negative evaluations of the review set that, in turn, reduce attitude certainty. In addition, the negative effect of cross-review incoherence on review set evaluations is weaker when the reviewers provide more contextual information behind their opinions. These findings deepen our understanding of inconsistency across multiple pieces of information; reveal the consequences of review evaluations during an understudied stage of consumers' decision-making process; and provide critical implications for review platforms, companies, and reviewers.

History: Paul Pavlou, Senior Editor; David (Jingjun) Xu, Associate Editor.

Supplemental Material: The online appendix is available at <https://doi.org/10.1287/isre.2022.1150>.

Keywords: online reviews • information incoherence • inconsistency • purchase deferral • helpfulness • credibility • online word of mouth • consumer decision making

Introduction

Online ratings and reviews are increasingly popular among consumers and critical for companies (Jabr et al. 2020). A distinctive feature of ratings and reviews is that they are typically inconsistent as it is rare for reviewers with diverse backgrounds and expectations to share the same opinions about a product or service¹ (Hu et al. 2017). Two separate literatures study the implications of inconsistent opinions among reviewers. Literature at the *product* level focuses on the dispersion of product ratings because most review platforms prominently display rating distributions in the form of bar charts (Sun 2012). A meta-analysis about the effect of rating dispersion on product sales reveals that greater rating heterogeneity generally

leads to lower sales (Babić Rosario et al. 2016). In contrast, literature at the *review* level focuses on inconsistency within a single review in the form of two-sided arguments and reveals its positive effect on prospective consumers' evaluations of the review's helpfulness and credibility (Schlosser 2011, Cheung et al. 2012).

Surprisingly, we know little about the connection between the outcomes of interest in these two literatures: evaluative judgments of reviews and product-level decision making. The prevailing assumption of helpful or credible reviews being persuasive may dissuade researchers from empirically examining the downstream consequences of review helpfulness or credibility. However, recent evidence from Yin et al. (2021) challenges this assumption and reveals an

exception in which favorably evaluated reviews may not be influential in affecting consumers' attitude and purchase decisions. Thus, the direct consequences and implications of consumers' evaluations of reviews are worthy of further investigation. If the helpfulness and credibility of reviews do not necessarily influence their persuasive power to change attitudes about a product or purchase decisions, then what could they influence more directly?

Moreover, why is inconsistency detrimental for sales at the product level but beneficial for evaluative judgments at the review level? Apart from the different outcomes, another likely reason is the scope of inconsistency at product versus review levels. The dispersion of ratings at the product level is derived from *all* product reviews, whereas a *single* two-sided review represents one reviewer's mixed opinions. Consumers seldom decide after reading a single review, but they rarely consult all reviews either. Instead, they routinely read multiple reviews in a short period of time to inform their decision making (Purnawirawan et al. 2012, Liu et al. 2019). In particular, the set of featured or top reviews (which we call a "review set") highlighted in nearly all review platforms are typically read by most consumers, and these reviews can influence how many more reviews they will read. Although two-sided reviews tend to be evaluated more favorably than one-sided reviews (Schlosser 2011, Cheung et al. 2012), little is known about how much this positive effect can extend to inconsistency across reviews at the level of a review set.

In this paper, we explore the following research question: how, why, and when would inconsistency *across top reviews* influence the *timing* of consumers' purchase decisions? Instead of focusing on purchase decisions or sales, our outcome of interest is purchase deferral: a key, yet understudied stage in consumers' decision-making process. After consumers read a number of reviews, they need to decide whether they are ready to make the purchase decision (i.e., to buy or not to buy) or defer this decision by "choosing not to choose" and seeking more information instead (Dhar and Nowlis 2004). Most consumers do not decide right away, especially if they are serious about a purchase (Krijnen et al. 2015). Distinct from the buy-or-not-buy *outcome* of the purchase decision that is the predominant focus of the e-commerce literature, purchase deferral is concerned with the timing of the purchase decision (Dhar 1996). A critical and often lengthy element of the purchase deferral stage is the prepurchase information search (Alba et al. 1997) for which online reviews provide a key resource. Consumers typically read the set of top reviews to guide their seeking of additional information (e.g., reviews) and gauge their readiness of making the final call. The set of reviews which consumers end up being exposed to and reading

can sway their preferences and purchase decisions (Liu and Karahanna 2017, Lei et al. 2022). Therefore, it is valuable to explore what characteristics of top reviews cause consumers to defer their decision and seek out more reviews.

We focus on cross-review incoherence: a unique form of inconsistency that cannot be observed within a single review or at the level of product ratings. Unlike two-sided arguments within a single review that usually provide pros and cons of *different* attributes of the same product, inconsistency among a set of reviews often (but not always) manifests as cross-review incoherence: a direct contradiction among different reviewers about the *same* attribute. To isolate this theoretical construct, we limit our attention to two particular forms of inconsistency among a set of reviews: *complementary* and *contradictory* inconsistencies. Both forms of inconsistency involve mixed opinions from different reviewers about the product—that is, some reviewers like the product, whereas others do not. However, in the case of contradictory inconsistency, there is disagreement at the attribute level, whereas in the case of complementary inconsistency, the disagreement only exists at the product level. We refer to the presence of attribute-level disagreement as *cross-review incoherence*.

Despite the prevalence of direct contradiction among real-world reviews, little research recognizes its importance or looks into its implications (cf. Liu and Karahanna 2017). We argue that the impact of cross-review incoherence on consumers' evaluative judgments of the review set is different from the positive impact of two-sidedness within a single review revealed in the prior literature. Based on cognitive dissonance theory, we hypothesize that cross-review incoherence reduces consumers' perceived helpfulness and credibility of the review set because of activated cognitive dissonance, and this negative effect can be reduced when specific contexts of different reviewers' opinions are made salient. We further argue that attitude certainty is key in explaining the consequences of review set evaluations. Specifically, we propose that lower evaluations of a review set's helpfulness and credibility can reduce the certainty of consumers' attitude toward the product, thus increasing their purchase deferral.

We conducted two experiments in which we manipulated the presence of cross-review incoherence and the boundary condition of context specificity. Our findings provide valuable insights beyond buy-or-not-buy decisions (by uncovering factors that drive purchase deferral), beyond a single review (by recognizing the need of consumers to read multiple but not all reviews), and beyond two-sidedness (by exploring the impact of cross-review incoherence, which does not exist at the review level). Our findings offer implications for review

platforms, companies, and reviewers. Because consumers' ultimate purchase decisions are influenced by the reviews that they actually read rather than all the available reviews (Liu et al. 2019, Lei et al. 2022), the knowledge of what drives purchase deferral under what circumstances can help companies estimate the likely number of reviews that would matter for most consumers, predict the timing of possible conversions or abandonments, and prioritize their efforts in dealing with a rapidly increasing number of reviews.

Theoretical Development and Hypotheses Inconsistency Within and Across Reviews

Inconsistency is common in online reviews. It manifests within an individual review as two-sidedness: presenting both positive and negative sides of a product (Jensen et al. 2013). Compared with one-sided reviews, two-sided reviews offer more diverse opinions, helping consumers to better evaluate a product's fit with their needs. Two-sided reviews are perceived as more helpful and credible than one-sided reviews (Schlosser 2011, Jensen et al. 2013). Yet consumers rarely base their purchase decision on a single review, and they typically consult multiple reviews. Whereas consumers might prefer two- to one-sided reviews, less is known about how they react to inconsistency in product opinions across multiple reviews. Conflicting opinions across reviews may offer some of the advantages of a single two-sided review, yet they may also create challenges that do not exist within a single review (Shan et al. 2021).

In this paper, we explore the effects of inconsistency across reviews on consumer decision making. To this end, we recognize that a product can be evaluated at two levels: the product as a whole and the product's attributes (i.e., features or functionalities) (Liu and Karahanna 2017).² Although evaluations of attributes often inform evaluations toward objects (products in our case) (Fishbein and Ajzen 1975), these two types of evaluations are distinct (Ledgerwood et al. 2018). For example, a consumer's preference for a product does not necessarily mean that the consumer likes all features of the product. Similarly, even if consumers dislike a product, they may still like certain features. Prior research in online word-of-mouth generally overlooks this important distinction (cf. Liu and Karahanna 2017).

The product-attribute distinction of evaluations has important implications for conceptualizing inconsistency within a single review and across multiple reviews. A two-sided review has positive opinions about some attributes and negative opinions about others (Schlosser 2011). However, two-sided reviews are expected to be coherent such that the positive and negative

opinions in the review refer to different attributes (Shan et al. 2021). For example, a reasonable review of a pair of over-ear headphones would not state that its battery life is both long and short. In sum, inconsistency in a two-sided review rarely occurs at the attribute level. In contrast, inconsistency across multiple reviews can exist at the product level *and* at the attribute level. Akin to two-sidedness in a single review, cross-review inconsistency occurs at the product level when the product receives mixed opinions: some reviewers like the product, whereas others dislike it. In addition, inconsistency across multiple reviews can also occur at the attribute level: in a review set, a specific attribute can be evaluated positively in one review and the same attribute can be evaluated negatively in another. We limit our attention to the presence of such attribute-level contradictions when product-level opinions are mixed, and we label it cross-review incoherence.

With a focus on cross-review incoherence (i.e., attribute-level contradictions) in a review set, we propose two types of inconsistency across multiple reviews: complementary inconsistency and contradictory inconsistency. This conceptualization builds on Morrison (2011), who states that inconsistency is complementary if it occurs in different parts of a complex system, and it is contradictory if it occurs in the same part of the system. In our context, if we consider a product as a system, then each product attribute is a part of that system. A review set has complementary inconsistency if multiple reviews in the set differ in their evaluations about the product but offer consistent evaluations regarding any specific attribute of the product. An example is a review set for headphones (product) with different opinions of the headphones overall (product-level inconsistency), but the reviews within the set maintain a consistent positive evaluation about the weight (attribute one) and a consistent negative evaluation about the battery life (attribute two). In contrast, a review set has contradictory inconsistency if multiple reviews in the set differ not only in their evaluations about the product, but also in their evaluations regarding the same attributes: some reviewers evaluate certain attributes positively, whereas other reviewers evaluate those same attributes negatively. In the headphones example, this occurs if, along with different opinions of the headphones overall (product-level inconsistency), some reviewers find them light (attribute one), whereas other reviewers find them heavy (attribute one). Table 1 compares the two types of inconsistency across reviews. Notably, although real-world review sets may contain both complementary and contradictory inconsistencies, we explicitly delineate these two concepts to isolate the theoretical construct of cross-review incoherence and theorize the impact of this construct on consumers' deferral of their buy-or-not-buy decisions.

Table 1. Distinct Types of Cross-Review Inconsistency in a Review Set

Inconsistency type	Inconsistency at product level?	Inconsistency at attribute level?	Explanation	Headphone example
Complementary	Yes	No	Inconsistency at the product level <i>only</i> . – Product evaluated differently across different reviews. – There is consistency at the attribute level: consistent positive or negative evaluation of an attribute across multiple reviews in the review set.	Overall, the headphones receive some positive and some negative reviews. – Some reviews have positive evaluations about the weight. – Other reviews have negative evaluations about the battery life.
Contradictory	Yes	Yes	Inconsistency at the product <i>and</i> attribute levels. – Product evaluated differently across different reviews. – There is disagreement across the reviews about an attribute: some reviews judge the attribute favorably and others judge it unfavorably.	Overall, the headphones receive some positive and some negative reviews. – Some reviews have positive evaluations about the weight. – Other reviews have negative evaluations about the weight.

Cross-Review Incoherence and Evaluative Judgments

We argue that cross-review incoherence affects consumers' perception of a review set's helpfulness and credibility. Helpfulness of an individual review is the degree to which the review is perceived to be useful in facilitating a consumer's decision-making process (Mudambi and Schuff 2010). Similarly, helpfulness of a review set is the degree to which the set is perceived to be useful (Purnawirawan et al. 2012). Review credibility refers to the extent to which the information in a single review is perceived to be believable or true (Tseng and Fogg 1999), and we define the credibility of a review set similarly.

We focus on these two evaluative judgments for three reasons. First, prior literature typically examines a review's helpfulness and credibility as outcome variables (Hong et al. 2017, Qahri-Saremi and Montazemi 2019), but these evaluations are also critical antecedents in a consumer's decision-making process (Cheung and Thadani 2012). Second, there is mixed evidence on the persuasive impact of helpful or credible reviews (Qahri-Saremi and Montazemi 2019, Yin et al. 2021). As we argue later, these judgments of reviews may influence decision deferral more directly than purchase decisions. Exploring this effect sheds light on the role of review helpfulness and credibility beyond the predominant focus on purchase decisions in the existing literature. Third, inconsistency within an individual review in the form of two-sidedness enhances the helpfulness and credibility evaluations of the review (Schlosser 2011, Cheung et al. 2012). Extending this investigation across reviews extends

our understanding of the distinct role of inconsistency beyond an individual review.

Based on the cognitive dissonance theory, we propose that cross-review incoherence negatively influences the perceived helpfulness and credibility of a review set. Cognitive dissonance refers to a psychological state of discomfort or stress triggered by factors such as contradictory information in the environment, inconsistency of one's belief with action or new information, etc. (Festinger 1962). Because individuals find it difficult to process self-contradictory information (Alter and Oppenheimer 2009), the presence of cross-review incoherence—direct contradiction in the evaluation of the same attributes among different reviews—increases cognitive dissonance.

Increased cognitive dissonance has a negative effect on perceived helpfulness of a review set. As people dislike being in a state of dissonance (Festinger 1962), they use different strategies to revert from this negative psychological state. One strategy is to depreciate the value of the contradictory information that is causing the dissonance (Darley and Gross 1983). Yin et al. (2016) observe such an effect when consumers evaluated an online review that activated cognitive dissonance as less helpful. We argue that this also holds for a review set when consumers experience dissonance among different reviews.

Increased cognitive dissonance also reduces the perceived credibility of a review set. Cognitive dissonance signals that something is not right, prompting people to be more critical and suspicious toward the information (Schwarz et al. 2016). In fact, the opposite of cognitive dissonance, fluency of information processing, is

one of the key mechanisms through which people judge credibility (Brashier and Marsh 2020). As an example, easy-to-read statements are perceived as more credible and believable than hard-to-read statements even when the statements are identical in their content (Reber and Schwarz 1999).

Taken together, we expect cross-review incoherence, which is present in the case of contradictory inconsistency but absent in the case of complementary inconsistency, to reduce both perceived helpfulness and perceived credibility of a review set.

Hypothesis 1. *The presence of cross-review incoherence in a review set has a negative impact on consumers' (a) perceived helpfulness and (b) perceived credibility of the review set.*

Context Specificity

Next, we focus on the moderating role of context specificity for the impact of cross-review incoherence on review set helpfulness and credibility. The evaluation of review content is context-dependent (Peng et al. 2020), and the knowledge of reviewers' context can buffer the cognitive dissonance caused by cross-review incoherence. A review's contextual information refers to user-related situations on which the opinions are based (Dridi et al. 2020). Building on the concept of information specificity (Choudhury and Sampler 1997), we define context specificity as the extent to which the opinions expressed in a review are accompanied by the specific context mentioned by the reviewer. If a reviewer describes the specific use case behind the reviewer's opinion of a certain attribute, context specificity is high because the opinion applies to only this reviewer's situation rather than to all situations.

As argued earlier, cross-review incoherence in review sets should lead consumers to experience more cognitive dissonance. However, this effect may be mitigated if the opposing opinions about the same attribute are accompanied by contextual information. When context specificity is low, opposing opinions from different reviews offer little insight into the specifics behind the opinions and, thus, appear to be more broadly applicable to most (if not all) consumers. As a result, consumers may have a harder time reconciling the conflicting opinions and are more likely to experience cognitive dissonance. In contrast, when context specificity is high, the articulation of specific user- or use-related situations limits the applicability of reviewers' opinions to a smaller set of consumers or specific circumstances. In this case, because different reviewers may have arrived at opposing opinions based on diverse situations, readers may self-select the review information that applies to their situation and, thus, are less likely to experience cognitive dissonance. In the headphones example, the review set with contradictory inconsistency

should cause more cognitive dissonance because it involves one review stating that the headphones are heavy and the other stating that they are light. However, if the first reviewer reveals that the reviewer uses them throughout the workday and the second reviewer reveals that the reviewer primarily uses the headphones for at most an hour at a time, the conflict between the two opinions and the resulting dissonance may be considerably reduced.

Thus, high context specificity in a review set should weaken the impact of cross-review incoherence on cognitive dissonance, which is negatively associated with evaluations of the review set.

Hypothesis 2. *Context specificity moderates the negative impact of cross-review incoherence on consumers' (a) perceived helpfulness and (b) perceived credibility of the review set such that this impact is lower when context specificity is high.*

Attitude Certainty and Purchase Deferral

Evaluations of a review set have critical implications on consumers' attitude certainty and purchase deferral. Attitude certainty refers to the degree of confidence with which one holds a particular attitude (Rucker et al. 2014). Attitude certainty is independent of attitude valence. We argue that the helpfulness of a review set positively influences attitude certainty. Compared with unhelpful reviews, helpful reviews have higher utility for the consumers' decision-making process (Liu and Zhang 2010). Greater utilization of helpful reviews enhances consumers' understanding of the product and leads to a better sense of whether it fits their needs (Hong and Pavlou 2014). Thus, review sets that are deemed more helpful should enable consumers to evaluate a product more confidently and increase the certainty of their attitudes.

Hypothesis 3 (a). *The helpfulness of a review set has a positive impact on attitude certainty.*

Credibility of a review set also positively influences attitude certainty. Untrustworthy information is typically discarded or discounted to avoid potential risks (Wathen and Burkell 2002). If a review set is not deemed credible, the consumer has not gained any additional information to increase the certainty of their judgment. However, if a review set is considered credible, the reviews contribute to the repository of information that the consumer can use to judge the product, resulting in increased certainty about that judgment (Tormala and Petty 2004). Thus, review sets with high credibility are more readily accepted and adopted into the consumer's decision-making process than review sets with low credibility.

Hypothesis 3 (b). *The credibility of a review set has a positive impact on attitude certainty.*

Attitude certainty, in turn, influences consumer purchase deferral, which is an important outcome of interest in consumer behavior research. Purchase deferral refers to postponing a decision to buy or not to buy a product (Tversky and Shafir 1992, Greenleaf and Lehmann 1995). Purchase deferral is distinct from purchase decisions. Deferral is concerned with decision timing that can be immediate or delayed, whereas the decision is concerned with the outcome that can be to buy or not to buy (Dhar 1996). Note that purchase deferral is not necessarily good or bad for consumers and businesses. Purchase deferral can be disadvantageous to consumers (e.g., increasing the cost of their invested time and effort) and to businesses (e.g., resulting in a financial loss if consumers decide not to buy after additional information search). Yet a search for additional information can also benefit consumers by leading to more informed and better buy-or-not-buy decisions. Further, this thoughtful process can prevent unrealistic expectations and result in less regret, reducing the likelihood of disappointed consumers to leave negative reviews.

A primary cause for purchase deferral is uncertainty in consumers' attitudes (Novemsky et al. 2007). When individuals are uncertain about their decision outcome, they tend to procrastinate and delay their decision (Solomon and Rothblum 1984, Greenleaf and Lehmann 1995). Also, in the face of uncertainty, people often continue their search of information to regain confidence over their decisions and to reduce perceived risk (Urbany et al. 1989, Shiu et al. 2011). Thus, reduced certainty in consumers' attitudes about a product makes it harder for them to make up their mind and delays their purchase decisions.

Hypothesis 4. *Attitude certainty regarding a product has a negative impact on purchase deferral (i.e., reducing the likelihood of consumers to defer the decision).*

Figure 1 depicts our theoretical model. We conducted two experiments in which we manipulated cross-review incoherence and context specificity. In

Study 1, we tested the direct effects of our manipulations on purchase deferral using a behavioral measure of the dependent variable. In Study 2, we extended the first study by testing the full conceptual model and by triangulating our findings with an alternative, intention-based measure of purchase deferral.

Study 1

In Study 1, we use a laboratory experiment to manipulate cross-review incoherence and context specificity in a 2×2 between-subjects design. We simulated a scenario in which participants were asked to read a set of four consumer reviews of a digital camera being sold on Amazon.com. We manipulated the two factors using different combinations of four individual reviews randomly selected from a review pool. After participants read a review set, they were given the opportunity to read additional reviews until they reported that they were ready to make a buy-or-not-buy decision.

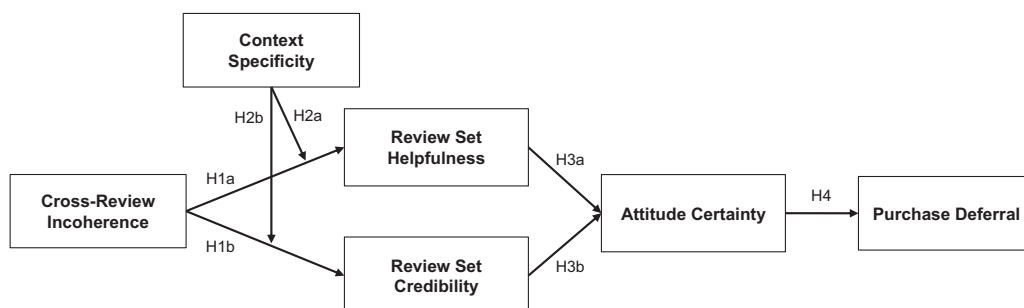
Sample

Participants ($n = 411$, 47.2% female) were recruited from an upper-level management class at a large, public university in the southern United States. Age ranged from 18 to 45 ($M = 19.0$, $SD = 2.42$).

Stimulus

The key stimuli in this experiment were review sets. Each review set consisted of four individual reviews that were adapted from the reviews used in Liu and Karahanna (2017) and actual Amazon reviews (see Table 1). Each individual review focused on one of two features: the autofocus function (feature A) or its user-friendliness (feature B), both of which are important attributes of digital cameras (Liu and Karahanna 2017). Although a review in reality can be two-sided (describing positive and negative aspects of the product), we designed each review as purely positive or negative to isolate the effects of cross-review incoherence, the central focus of our research. Isolating these effects required us to keep individual reviews relatively

Figure 1. Research Model



simple—having only one valence and addressing only one feature. More complex reviews (e.g., two-sided reviews) would introduce extraneous variability not directly relevant to the study. Specifically, we first wrote two positive reviews for each feature (A1+, A2+, B1+, B2+). Reviews A1+ and A2+ were worded very differently to appear as being written by different reviewers. The same applied to reviews B1+ and B2+. Then, negative versions of each review were written (A1−, A2−, B1−, B2−). Substantive content was kept as similar as possible between the positive and negative versions of each review to remove possible confounds.

Using a pool of these eight reviews, we could vary the presence of cross-review incoherence. In each of the review sets that we constructed, two reviews were negative and two were positive. In the review sets without cross-review incoherence (i.e., complementary inconsistency), one camera feature exclusively received negative reviews and the other feature exclusively received positive reviews (i.e., A1−, A2−, B1+, B2+ or A1+, A2+, B1−, B2−). In the review sets with cross-review incoherence (i.e., contradictory inconsistency), each camera feature was rated positively by one reviewer and negatively by another. Thus, the reviews directly contradicted each other for each of the two features (i.e., A1+, A2−, B1+, B2− or A1−, A2+, B1−, B2+).

We manipulated context specificity by including more concrete details about user- or use-related situations (Schwanenflugel et al. 1988) (shown in brackets in Table 2) in the *high context specificity* conditions and by excluding this content in the *low context specificity* conditions. Contextual information was as similar as possible (other than valence) between the positive and negative versions of each review.

To operationalize our dependent variable—a behavioral measure of purchase deferral—we created 32 additional individual reviews, 22 of which were positive and 10 of which were negative. These additional reviews were adapted from the same source as the reviews in the main set. We crafted these reviews to make them generic without getting into specific features (especially the two features used in the main review sets) and to keep them similar in length and affective intensity (see Online Appendix A). By doing so, we avoided the possibility that the additional reviews would operate differently across conditions. For example, for participants in the conditions with contradicting reviews about the autofocus feature, any additional review that mentioned autofocus may influence these participants differently from others who read contradicting reviews about the camera’s user-friendliness feature. Because participants in the former case may look for autofocus-related reviews,

Table 2. Content of Individual Reviews for Study 1

	Positive reviews (+)	Negative reviews (−)
A1	The autofocus function is very good. It is very fast and accurate [in bright light situations]. I have been able to shoot decent pictures of my children [standing still] using the autofocus function. [While I have only taken pictures during daytime,] I did not find any blurry parts in the images at all.	The autofocus function is very bad. It is very slow and inaccurate [in low light situations]. I have been unable to shoot decent pictures of my children [running around] using the autofocus function. [While I have only taken pictures during nighttime,] I found many blurry parts in the images.
A2	The auto focus of this camera is very useful, and I am impressed. Because it’s so quick, the picture is rarely out of focus [when light is adequate]. Just yesterday [after lunch], I took several pictures of my friends. Those pictures came out sharp and clear[, but I have steady hands].	The auto focus of this camera is almost useless, and I am unimpressed. Because it’s so slow, the picture is often out of focus [when light is inadequate]. Just yesterday [after dinner], I took several pictures of my friends. Those pictures came out fuzzy and unclear[, but my hands sometimes shake].
B1	This camera is very user-friendly. [After a few hours of use,] I find it really straightforward. [Although I had to try it a few times,] I feel it is easy to figure out the various settings/functions. I also like how intuitive the buttons are[; admittedly, I spent a lot of time getting to know the camera].	This camera is not user-friendly at all. [After a few minutes of use,] I find it really complicated. [Although I tried it only once,] I feel it is difficult to figure out the various settings/functions. I also dislike how counterintuitive the buttons are[; admittedly, I spent very little time getting to know the camera].
B2	I find the camera quite easy to operate. Finding the right buttons to perform various functions isn’t a problem at all [as long as you are familiar with these cameras]. [Though I have small hands, for me] the buttons and controls are well designed. Switching between different settings (e.g., portrait mode, landscape mode) is also quick and easy [if you have read the manual].	I find the camera quite difficult to operate. Finding the right buttons to perform various functions is a big problem [if you are unfamiliar with cameras]. [Though I have big hands, for me] the buttons and controls are poorly designed. Switching between different settings (e.g., portrait mode, landscape mode) is also slow and hard [if you have not read the manual].

Notes. A = reviews about the autofocus feature; B = reviews about the user-friendliness feature; 1 and 2 = different reviews about the same feature. The content in brackets was only shown to participants in the high context specificity conditions.

incorporating opinions about autofocus into the additional reviews may create confounds that may covary with our manipulations. It may also complicate and possibly contaminate our measure of purchase deferral (i.e., the number of additional reviews that participants want to read after reading the treatment reviews): “earlier” additional reviews may influence participants’ willingness to read “later” additional reviews differently across conditions. We address this concern more fully in Study 2 by using an intention measure of the outcome variable.

Procedure and Measures

Participants were asked to evaluate a digital camera available on Amazon.com, referred to as “Model X.” Participants were informed that they could read information about the camera before deciding whether to purchase it. After seeing a screenshot from Amazon depicting the camera and its features (see Online Appendix B), they were asked to read a set of four “verified purchase” reviews that were ostensibly the most recently posted. These reviews were used to create four experimental conditions in a 2×2 between-subjects design to which participants were randomly assigned. In addition to manipulating the two factors (cross-review incoherence and context specificity), we also counterbalanced which feature would get positive (or negative) reviews to account for preferences that participants may have had for one feature over the other. Specifically, half of the participants were randomly assigned to view negative reviews on the autofocus feature and positive reviews on the user-friendliness feature (A1–, A2–, B1+, B2+), whereas the other half viewed positive reviews on the autofocus feature and negative reviews on the user-friendliness feature (A1+, A2+, B1–, B2–). Online Appendix C provides two example review sets.

After reading a set of four reviews, participants were asked to respond to the following question: “If you were thinking of buying a digital camera, how ready are you to decide whether to purchase Model X right now?” The response options were (a) “I am *not ready* to make this decision; I would like to read more reviews”; (b) “I am *almost ready* to make this decision; I would like to read one or two more reviews”; and (c) “I am *ready* to make this decision.” If participants selected one of the first two options, they were shown an additional product review, and then the question was repeated. Participants could read up to 32 additional reviews.³ To ensure the internal validity of the experiment design, we used a fixed sequence for the additional reviews in all conditions with two positive reviews followed by a negative review. We chose to overrepresent positive reviews because this is more consistent with the distribution of reviews in reality (Chevalier and Mayzlin 2006). Participants could not

choose which review they would see next. The dependent variable was calculated as the number of additional reviews participants read until they indicated being ready for a purchase-or-not decision. This experimental paradigm (i.e., repeated iterations of acquiring new information and making a decision) is used in the psychological literature to test attitudinal changes over time as new information is acquired (e.g., Haselhuhn et al. 2010).

Manipulation Checks

We recruited 160 participants from Amazon’s Mechanical Turk to pretest the effectiveness of these manipulations. Two manipulation checks were used for the cross-review incoherence manipulation at the review set level. Using a seven-point Likert scale (1 = strongly disagree, 7 = strongly agree), participants evaluated the review set as a whole. First, participants responded to a three-item measure of perceived contradiction we developed based on Ahn et al. (2011). An example item is “Some reviewers expressed conflicting opinions about the same aspect of the product.” Cronbach’s alpha was $\alpha = 0.93$. As expected, perceived contradiction was significantly higher, $t(158) = 9.64, p < 0.001$, in the incoherence-present conditions ($M = 5.80, SD = 1.13$) than in incoherence-absent conditions ($M = 3.70, SD = 1.60$). Second, participants responded to a three-item measure of perceived consistency (Cheung et al. 2012). An example item is “The opinions expressed in these reviews are consistent with each other.” Cronbach’s alpha was $\alpha = 0.94$. Again, as expected, perceived consistency was significantly lower, $t(158) = -4.32, p < 0.001$, in the incoherence-present conditions ($M = 2.71, SD = 1.27$) than in the incoherence-absent conditions ($M = 3.61, SD = 1.35$).

To evaluate the contextual specificity manipulation, participants responded to a five-item measure of perceived contextual specificity that we developed based on Goodman et al. (2004). Using a seven-point Likert scale, participants evaluated each individual review. An example item is “The reviewer described how the product worked in a specific context.” The average Cronbach’s alpha across the ratings of the four reviews was $\alpha = 0.89$. As expected, perceived contextual specificity was significantly higher, $t(158) = 2.81, p = 0.006$, in the high context specificity conditions ($M = 5.32, SD = 0.75$) than in the low context specificity conditions ($M = 4.98, SD = 0.79$).

Results and Discussion

We first examined if having a finite number of reviews available to read resulted in a ceiling effect. The range for this variable was 0 to 31 ($M = 2.27, SD = 3.20$), indicating no ceiling effect. Online Appendix E provides frequency statistics. We then examined the effects of cross-review incoherence and context specificity on

purchase deferral. As our behavioral measure (the number of additional reviews read) is a count variable, we first tested if this variable followed a Poisson distribution. The Kolmogorov–Smirnov Z-test was significant ($p < 0.001$), indicating that this measure did not follow a Poisson distribution and was over-dispersed. Thus, we used a negative binomial model to analyze our data and controlled for the order of reviews in the set. The main effect of cross-review incoherence was marginally significant, Wald χ^2 ($df = 1$) = 3.72, $p = 0.054$. Participants tended to read more reviews in the incoherence-present conditions ($M = 2.38$, $SE = 0.20$) than in the incoherence-absent conditions ($M = 1.91$, $SE = 0.16$). The main effect of contextual specificity was nonsignificant, Wald χ^2 ($df = 1$) = 0.96, $p = 0.326$. The interaction between cross-review incoherence and contextual specificity was significant, Wald χ^2 ($df = 1$) = 6.06, $p = 0.014$. An analysis of the marginal effects showed that, when context specificity was low, participants read significantly more reviews in the incoherence-present conditions compared with the incoherence-absent conditions, $t(203) = 2.83$, $p = 0.005$, whereas the difference was nonsignificant when context specificity was high, $t(204) = -0.36$, $p = 0.720$ (Figure 2). Together, these results support the overall effect of cross-review incoherence and its interaction with context specificity implied by our theoretical framework. In Study 2, we extend these findings by examining the full conceptual model. We also triangulate the results by using an alternative, intention-based measure of the outcome variable.

Study 2

Study 2 investigates the hypothesized mediators, again using a laboratory experiment to manipulate cross-review incoherence and context specificity in a 2×2 between-subjects design. We use the same scenario and stimuli as in Study 1. The procedure differs in that, after participants read a review set, they completed measures of the mediating variables and their intention to defer their buy-or-not-buy decisions.

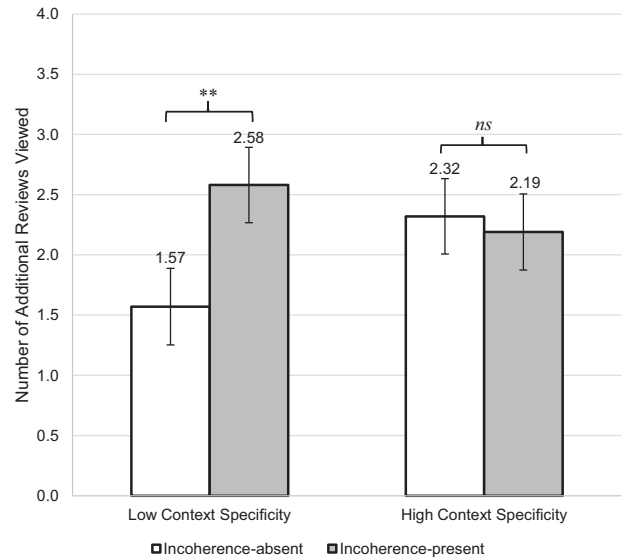
Sample

Participants ($n = 737$, 51.4% female) were recruited from an upper-level management class at a large, public university in the southern United States. Age ranged from 18 to 64 ($M = 21.0$, $SD = 4.04$).

Stimulus and Procedure

We used the same review sets from Study 1 to create the same four conditions. After reading the reviews, participants evaluated the review set in terms of helpfulness and credibility. Participants then reported their attitude about the product, their certainty about their attitude about the product, and whether they

Figure 2. Effects of Cross-Review Incoherence and Context Specificity on the Number of Additional Reviews Viewed



Notes. ± 1 standard error bars are shown. ns, not significant.
 ** $p < 0.01$.

were ready to make a final purchase decision.⁴ Finally, they answered two manipulation check questions.

All measures except purchase deferral used a seven-point scale (see Online Appendix F for all measurement items). Review set helpfulness was measured with three items on a semantic differential scale (Yin et al. 2021). For example, participants were asked to describe the set of reviews (as a whole) on a scale from “not at all helpful” to “very helpful.” Cronbach’s alpha was $\alpha = 0.85$. Perceived credibility was measured with four items on a semantic differential scale adapted from Cheung et al. (2012). For example, participants were asked how they would describe the set of reviews from “not at all credible” to “very credible.” Cronbach’s alpha was $\alpha = 0.86$. Attitude certainty was measured with three Likert items on a scale from one (not at all) to seven (very) (DeMotta et al. 2016). One example item is “How confident are you feeling that you made the right judgment about this camera?” Cronbach’s alpha was $\alpha = 0.92$. We examined the convergent and discriminant validity of the three mediating constructs by conducting an exploratory factor analysis (see Online Appendix G for details).

Purchase deferral was measured with a single question (Pang et al. 2017): “Please choose among the following options to indicate the statement that is closest to your final decision about Model X.” Four options were provided: (a) “I would have chosen to purchase this model,” (b) “I would have chosen to not purchase this model,” (c) “I would have chosen to search for more reviews of Model X and decide later,” and (d) “I

Table 3. Descriptive Statistics and Correlations Between All Variables in Study 2

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6
1. Cross-review incoherence	1.49	0.50	–					
2. Context specificity	1.50	0.50	0.00	–				
3. Review set helpfulness	5.10	1.25	–0.19**	0.11**	0.85			
4. Review set credibility	4.65	1.01	–0.22**	0.06 [†]	0.52**	0.85		
5. Attitude certainty	4.75	1.37	–0.08*	0.03	0.36**	0.42**	0.92	
6. Purchase deferral	0.81	0.40	–0.07 [†]	–0.00	–0.11**	–0.16**	–0.18**	–

Notes. *N* = 737. Based on participants' assigned conditions, the manipulation of cross-review incoherence was coded as 1 = absent and 2 = present. Similarly, the manipulation of context specificity was coded as 1 = low and 2 = high. Cronbach's alphas are listed in the diagonal.

[†]*p* < 0.10; **p* < 0.05; ***p* < 0.01.

would have chosen to search for other models and decide later." Options (a) and (b) were coded as zero (no deferral), and options (c) and (d) were coded as one (deferral).

Results

Descriptive statistics and correlations between all variables are shown in Table 3. First, we tested the overall effect of cross-review incoherence. Given that our manipulated factors and dependent variable were all dichotomous, we used a chi-square test. The effect of cross-review incoherence on purchase deferral was marginally significant, $\chi^2(1) = 3.32$, *p* = 0.069. The effect of context specificity on purchase deferral was not statistically significant, $\chi^2(1) = 0.01$, *p* = 0.928. Frequency statistics are available in Online Appendix H.

Second, we examined the effects of our manipulation on the first stage of mediators, review set helpfulness and perceived credibility, using a two-way ANOVA. Statistically significant differences in both variables were observed between the incoherence-present and -absent conditions. Specifically, the review set was perceived to be less helpful in incoherence-present conditions (*M* = 4.86, *SD* = 1.37) than in incoherence-absent conditions (*M* = 5.33, *SD* = 1.07), *t*(735) = –5.19, *p* < 0.001, Cohen's *d* = –0.38. In addition, the review set was perceived as less credible in the incoherence-present conditions (*M* = 4.42, *SD* = 1.05) relative to the incoherence-absent conditions (*M* = 4.87, *SD* = 0.93), and this difference was significant, *t*(735) = –5.73, *p* < 0.001, Cohen's *d* = –0.45. These results provide support for Hypotheses 1(a) and 1(b). Context specificity also had a significant positive effect on review set helpfulness, *t*(735) = 3.09, *p* = 0.002, Cohen's *d* = 0.23, and a significant positive effect on perceived credibility, *t*(735) = 2.36, *p* = 0.018, Cohen's *d* = 0.17. Although the main effects of contextual specificity are outside our scope, they are in the expected direction.⁵

Third, we examined whether the effects of cross-review incoherence were moderated by context specificity. This interaction was statistically significant for both review set helpfulness, *F*(1, 733) = 6.40, *p* = 0.012, and review set credibility, *F*(1, 733) = 9.33, *p* = 0.002 (see Figure 3). Pairwise comparisons show that, when context specificity was low, participants in

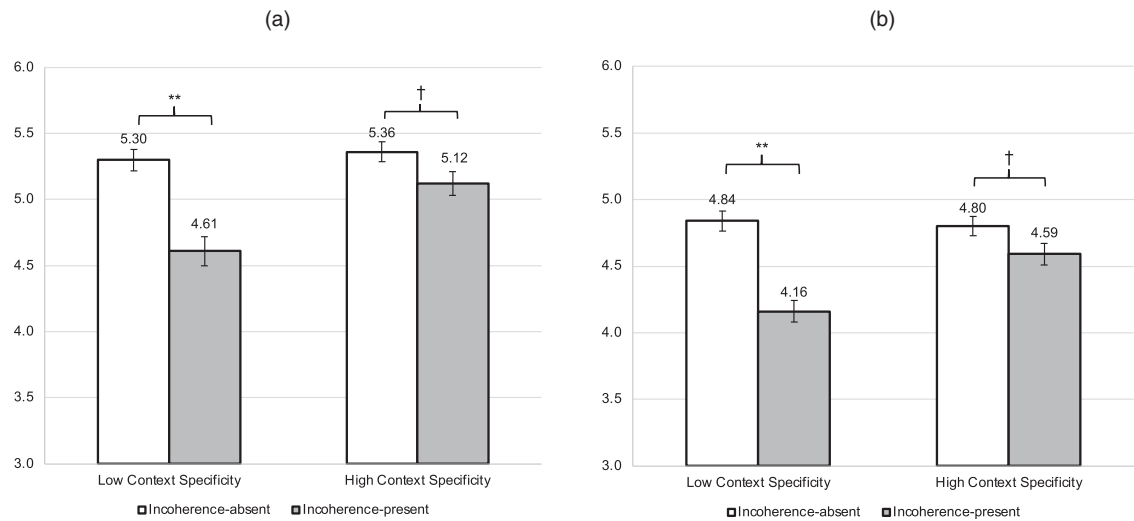
the incoherence-present conditions rated the review set as significantly less helpful, *t*(369) = –5.42, *p* < 0.001, and less credible, *t*(369) = –6.01, *p* < 0.001, than those in the incoherence-absent conditions. When context specificity was high, the difference across incoherence conditions was substantially smaller and marginally significant for review set helpfulness, *t*(384) = –1.89, *p* = 0.060, and credibility, *t*(384) = –1.86, *p* = 0.064. These results provide support for Hypotheses 2(a) and 2(b).

Finally, we tested our full moderated mediation model using the PROCESS macro that supports a dichotomous dependent variable (Hayes 2018). We built a custom model for testing serial moderated mediation (Hayes 2015) and used 5,000 iterations of bootstrapping to construct a bias-corrected confidence interval around the estimates (Hayes 2018). The path from attitude certainty to purchase deferral reflects a log-odds metric because the dependent variable is dichotomous. As shown in Figure 4, cross-review incoherence was significantly related to both review set helpfulness and credibility. Context specificity also significantly moderated these relationships. The effects of review set helpfulness and credibility were significantly and positively related to attitude certainty, which was, in turn, significantly and negatively related to purchase deferral. Thus, Hypotheses 3(a), 3(b), and 4 are all supported. The indirect effects of cross-review incoherence on purchase deferral via both pathways (i.e., review set helpfulness and credibility) were significant when context specificity was low and when it was high. However, the indirect effects were significantly stronger when context specificity was low as indicated by pairwise contrasts of the indirect effects (see Table 4).

Discussion

The results of Study 2 show that cross-review incoherence in a review set significantly decreased review set helpfulness and credibility, which, in turn, decreased certainty in participants' attitudes about a product. As a result, participants were more likely to defer their purchase decision. This effect was significantly reduced when context specificity was high. The results of Study 2 support all hypotheses.

Figure 3. Effects of Cross-Review Incoherence and Context Specificity on (a) Review Set Helpfulness and (b) Review Set Credibility



Note. ± 1 standard error bars are shown.
 $^{\dagger}p < 0.10$; $*p < 0.05$; $**p < 0.01$.

General Discussion

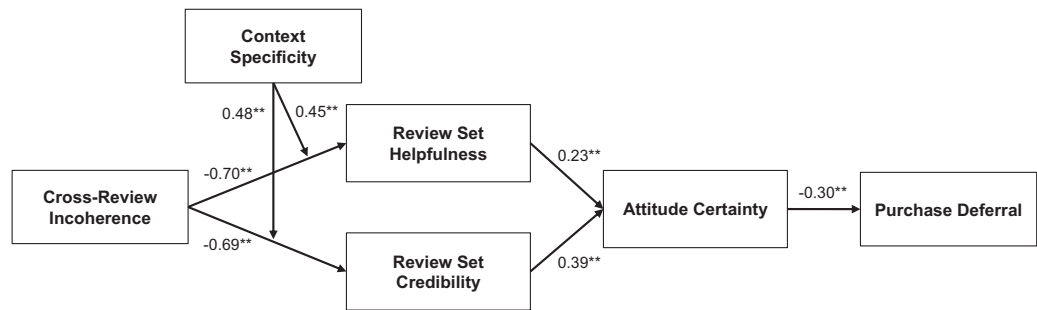
In two laboratory experiments, we manipulated the presence of cross-review incoherence and context specificity, and we measured participants’ purchase deferral through their behaviors or intentions. We found consistent evidence that (a) the presence of cross-review incoherence results in a higher likelihood of purchase deferral, (b) this effect occurs through helpfulness and credibility perceptions of the review set and then attitude certainty, and (c) the effect is reduced or nonexistent when context specificity is high. These findings provide important and useful implications for both theory and practice.

Theoretical Implications

First, this paper contributes to the online review literature by examining drivers of purchase deferral, an

important step in consumers’ decision-making process. Prior studies focus primarily on the impact of rating characteristics (such as rating valence and volume) on product sales based on observational data (Floyd et al. 2014, Babić Rosario et al. 2016). However, recent empirical evidence suggests that consumers read some (but not all) reviews before making buy-or-not-buy decisions (Liu et al. 2019). Moreover, to help consumers navigate the large quantity of reviews, platforms typically display a small subset of the most helpful reviews. After reading the highlighted reviews, consumers need to decide if they are ready to make the buy-or-not-buy decision or if they will defer and seek more information. This research is among the first to explore the drivers of consumers’ tendencies to defer their purchase decision after reading a series of reviews. Our findings shed light on the factors that influence the timing of buy-or-not-

Figure 4. Path Analysis Results (Study 2)



Notes. Unstandardized path estimates are shown. For clarity, the following paths were omitted from the figure: cross-review incoherence → attitude certainty ($b = 0.07$, $p = 0.48$), cross-review incoherence → purchase deferral ($b = 0.21$, $p = 0.28$), review set helpfulness → purchase deferral ($b = -0.03$, $p = 0.78$), review set credibility → purchase deferral ($b = -0.19$, $p = 0.07$).
 $*p < 0.05$; $**p < 0.01$.

Table 4. Indirect Effects and Pairwise Contrasts

Paths	Indirect effect (SE)	95% confidence interval
Incoherence → helpfulness → certainty → deferral		
Moderator: low context specificity	0.048** (0.02)	[0.017, 0.096]
Moderator: high context specificity	0.017* (0.01)	[0.001, 0.043]
Index of moderated mediation: −0.041, 95% CI = [−0.07, −0.01]		
Incoherence → credibility → certainty → deferral		
Moderator: low context specificity	0.079** (0.31)	[0.031, 0.146]
Moderator: high context specificity	0.024 [†] (0.02)	[−0.001, 0.063]
Index of moderated mediation: −0.054, 95% CI = [−0.11, −0.01]		

Notes. SE = standard error. 95% confidence intervals are bias-corrected.

[†] $p < 0.10$; * $p < 0.05$; ** $p < 0.01$.

buy decisions and open up a new area of inquiry into consumers' decision-making process.

Second, this work goes beyond a single review and focuses on the level of review sets, which is largely overlooked by past research on reviews. In addition to product-level studies examining the impact of ratings on sales, a growing literature focuses on individual reviews and a variety of factors that influence consumer perceptions of review helpfulness and credibility (e.g., Hong et al. 2017). Review helpfulness and credibility are popular outcome variables as it is assumed that a helpful or credible review is more persuasive in driving attitude and purchase decisions (Benlian et al. 2012). In reality, however, consumers rarely read a single review, and they routinely consult multiple reviews (Purnawirawan et al. 2012). A few recent studies address this middle level and explore product-level consequences of a set of top reviews. For example, Lei et al. (2022) show that a few top reviews can sway consumers' purchase decisions from the influence of product-level average ratings, which are commonly believed to be the primary determinant of product sales. In addition, the experimental evidence of Yin et al. (2021) shows that helpful reviews are not necessarily persuasive. If review helpfulness and credibility do not necessarily determine attitude and purchase decisions, then what is their role in consumers' decision making? Our paper complements these studies by exploring how consumers make helpfulness and credibility judgments of a set of reviews and by extending the ultimate outcome variable of interest from purchase decisions to *purchase deferral*. We propose that the direct consequences of review helpfulness and credibility are certainty about attitudes and deferral of buy-or-not-buy decisions. Our findings highlight the relevance of information search to evaluative judgments of helpfulness and credibility, and they suggest the need for future research to study the

consequences of evaluative judgments in addition to their antecedents.

Third, this paper deepens our understanding of information inconsistency and emphasizes the importance of studying cross-review incoherence that does not exist within a single review. Inconsistency in a single review typically manifests as two-sided arguments, which are found to have a *boosting* effect on perceived review helpfulness and credibility (Schlosser 2011, Cheung et al. 2012). In contrast, we propose and find a *dampening* effect of greater inconsistency among reviews at the level of a review set, suggesting that insights from prior studies on what constitutes a helpful review cannot be directly extended and applied to a set of reviews (see also Purnawirawan et al. 2012). As such, our work complements and extends recent studies on the role of conflicts among a set of reviews in consumer decision making (e.g., Liu and Karahanna 2017). Just as organizational scholars are cautioned to generalize individual-level phenomena to the group level (Klein and Kozlowski 2000), so too should we take caution in assuming that effects observed regarding individual reviews generalize to sets of reviews.

Moreover, we introduce a new type of inconsistency: cross-review incoherence. Two-sided opinions expressed within a single review rarely involve direct contradiction because people want to hold and express opinions that are consistent with each other (Festinger 1962, Elliot and Devine 1994). If a reviewer presents mixed opinions about a product, those opinions are typically expressed toward different attributes rather than toward the same attribute. As a result, direct contradiction receives little attention in review helpfulness research. When opinions come from multiple reviewers, however, inconsistency can manifest as direct contradictions on the same attribute (Liu and Karahanna 2017), which can further influence consumers' cognitive dissonance. Our studies demonstrate that

different types of inconsistency should not be treated equally and the detrimental impact of cross-review incoherence is worth future exploration. By providing indirect evidence for the critical role of cognitive dissonance, this work also complements and extends recent research that explores how consumers integrate multiple and oftentimes inconsistent pieces of information in decision making (e.g., Purnawirawan et al. 2012, Qiu et al. 2012, Yin et al. 2016, Shoham et al. 2017).

Fourth, we examine a critical contingency for the influence of cross-review incoherence: the specificity of contextual information. The influence of content on consumers is rarely context-free. Consumers' evaluation of user-generated content may depend on contextual information that accompanies it (Peng et al. 2020). Whereas the influence of cross-review incoherence on purchase deferral occurs through lower perceived review helpfulness and credibility, we demonstrate that these mechanisms can be weakened if reviews include specific contexts behind the opinions. Thus, this research complements recent findings that consumers' evaluation of online reviews is contingent on contextual factors (e.g., Yin et al. 2016) and that content-context interactions should receive more attention in future research.

Finally, our findings may offer broader theoretical implications beyond the setting of online review assessments. Our primary focus relates to the role of cognitive dissonance in individual decision making, in which an individual uses multiple pieces of conflicting information to determine whether and how much they need additional information. Thus, our findings may apply in other settings in which decision makers rely on a collection of (often conflicting) opinions, such as editors and program officers working with review panels for publications and funding proposals and media panels offering commentary to inform the public. Our findings may also be relevant in different IS- or IT-related settings, such as CIOs deciding on which enterprise software application to acquire based on diverse opinions from different stakeholders; individuals deciding whether to contribute toward a crowd-funding effort based on other donors' testimonies; and AI designers developing decision models that process multiple, contradictory pieces of information.

Practical Implications

Our findings also offer implications for review platforms, companies, and reviewers. Platforms almost universally sort reviews based on their helpfulness votes and then highlight the few most helpful ones. The number of reviews to display prominently is a key decision for review platforms; for example, Amazon.com displays three to six top reviews. Our findings suggest that some design choices of displaying reviews—such as the number of reviews to highlight—could benefit from an analysis

of how the top reviews' content is similar or different from each other. In particular, review platforms might analyze the interrelationships among a set of reviews and the specificity of use-related context using text-mining techniques to predict the likelihood and the extent of consumers to seek and read more reviews (Zhang et al. 2022). If three top reviews are unlikely to drive a quick decision (e.g., when reviews contain direct contradictions about the same feature), then it may be useful to display more reviews in anticipation of consumers' tendency to read more reviews. Another design choice is to automatically extract parts of reviews that address each product attribute, create a summary assessment for the attribute, and then display these aggregated attribute scores on the product page. This may help consumers to grasp the crowd's overall assessment of each attribute and reduce their struggle in reconciling contradictory opinions regarding an attribute.

Whereas we focus on the timing of purchase decisions, our study can also inform retailers and product manufacturers on how additional reviews influence buy-or-not-buy decisions. The timing of buy-or-not-buy decisions is important not because of the absolute length of the delay, but because it can influence the number and type of additional reviews to which consumers are exposed. Recent empirical evidence suggests that the set of reviews which consumers are exposed to and end up reading directly impacts preferences and purchase decisions (Liu and Karahanna 2017, Liu et al. 2019). Moreover, just a few reviews can sway consumers' ultimate decisions from the presumably dominant influence of a product's overall average rating, which is an all-encompassing signal of product quality (Lei et al. 2022). Even if consumers spend just a few more minutes reading some additional reviews, these additional reviews may change consumers' buy-or-not-buy decision. Thus, the timing of such decisions has significant implications for the platform and various parties benefiting from the platform. For example, retailers and manufacturers can use our framework to anticipate the likelihood of consumers to find the set of top reviews sufficient for a purchase decision. If companies expect the displayed review set to be unhelpful or not credible despite each review individually being very helpful (e.g., with contradictory opinions across reviews), they can make efforts to address the inconsistencies or apparent contradictions, such as offering a response or providing some context to mitigate purchase deferral. In addition, if the displayed review set is not deemed helpful or credible, consumers will consult more reviews. In this case, companies should not limit their attention to only those top reviews, but expand their focus to more reviews, such as the most recent reviews that consumers may consult next. Thus, our findings provide more nuanced, evidence-based guidance for companies to effectively manage and

respond to a rapidly increasing number of product reviews.

Our findings also benefit reviewers in crafting more helpful content. Although being selected as one of the “most helpful reviews” is a worthy achievement, reviewers should realize that consumers rarely make purchase decisions based on a single review. Thus, reviewers should consider opinions from other reviewers to truly help prospective consumers. When a reviewer holds contradicting opinions on an attribute of the product compared with other reviewers, then the reviewer should provide sufficient and detailed background information to justify the opinions and enhance context specificity. This may not only boost an individual review’s perceived helpfulness (e.g., Mudambi and Schuff 2010), but also reduce the likely impact of cross-review incoherence on consumers’ purchase deferral.

Limitations and Future Research

Whereas this paper facilitates a deeper understanding of the role of cross-review incoherence in consumer purchases, several limitations of our studies offer exciting opportunities for future research. First, although laboratory experiments are the gold standard to isolate the causal impact of theoretical constructs of interest in a phenomenon, they rely on a constructed simplification of reality to facilitate researcher control. The design of our experiments differs from reality in several aspects that are worthy of additional investigation. For example, our experiments lack monetary consequences as participants did not have to buy the product. We believe this to be a limited concern as our experiments represent a more conservative test of the hypotheses; the effects are expected to be even larger in real-world situations in which economic consequences are present. Our participants were given a fixed set of reviews instead of being allowed to form their own set of “featured reviews” or change their sequence. Also, after reading the top reviews, buyers may engage in other actions to gather information besides reading more reviews, such as visiting other websites. Our participants did not have that choice. We kept these factors constant in our experiments to achieve high control (i.e., isolating the effects of variables of our interest) and interval validity. This came with the natural cost of some mundane realism or external validity (Aronson et al. 1998). Therefore, future research should test, replicate, and extend our findings in settings in which these constraints can be relaxed and effectively addressed using complementary methods, such as randomized field experiments and controlled experiments with financial consequences.

Second, we demonstrate that the presence of *direct contradiction* among reviews at the *attribute* level

reduced perceived credibility of the review set. This extends prior research on individual reviews: inconsistency of a review with aggregated rating statistics leads to lower perceptions of the review’s credibility and diagnosticity (Qiu et al. 2012, Yin et al. 2016). Yet conflicting opinions are a defining characteristic of product reviews, and such inconsistencies might arise from variations in product quality, differences in reviewers’ expectations, or how they use the product. In contrast, a homogeneous set of reviews without conflicting opinions might look suspicious, and it is possible for consumers to believe such reviews to be “incentivized.” Although a comparison of the impacts of consistent and inconsistent reviews is outside the scope of this study, it is a fascinating direction to explore in future research.

Third, as we focused on cross-review incoherence, we developed the stimuli to match its definition and cleanly manipulate this variable. Specifically, we designed a collection of stylized reviews to ensure internal validity and make causal inferences, which are common practice in prior research (e.g., Yin et al. 2014, 2017; Lei et al. 2021). Specifically, the purpose of our review design was to control for confounds resulting from the product aspects being discussed and the linguistic style in which the reviews were written. The resulting reviews only differed in terms of the presence of cross-review incoherence and the specificity of contextual information but not in terms of substantive content or tone to reduce any threats to internal validity of our findings. Still, we acknowledge that reviews on commercial platforms are not always comparable with those in our experiments as they demonstrate more variety in terms of valence, length, breadth, tone, and language use. On one hand, real-world review sets may be consistent among different reviews, and future research can explore how inconsistent reviews would shape consumer behavior differently from consistent reviews (e.g., the effects of contradictory inconsistency should be greater as compared with the absence of product-level inconsistency vs. complementary inconsistency). On the other hand, the valence of all attribute evaluations need not align with the overall evaluation, and some reviews are two-sided. Although such reviews do occur in reality (but less frequently than the one-sided reviews that we used), they cannot help isolate the impact of cross-review incoherence or test our proposed relationships. Nevertheless, review two-sidedness has nontrivial impacts on consumers’ evaluative judgments of the reviews (Schlosser 2011, Lei et al. 2021). Looking beyond an individual review, future research should explore how consumers make sense of a combination of one- and two-sided reviews in a review set in their decision-making process.

Fourth, because we sought to examine if context specificity would moderate the effects of cross-review incoherence, the direct impacts of this moderator on the outcome variable and mediators were outside our scope. Contextual information may be of different types (e.g., temporal cues, location), and research on these types and their main effects represent a fascinating opportunity. Also, context can mean different things for different people. Consumers may differ in their likelihood to filter reviews or change their orders to fit their needs. Some consumers may delay the decision by several minutes, but others may spend hours reading more reviews, especially if the product is expensive or important (Krijnen et al. 2015). The exploration of consumers' individual differences is a worthy direction for future research.

Fifth, although we argue that cross-review incoherence generally reduces review set helpfulness and credibility, this effect is unlikely to be universal. Whereas we demonstrate one boundary condition—the specificity of contextual information—there may be other boundary conditions that can strengthen or weaken the impact of cross-review incoherence. For example, the negative impact of direct contradiction on helpfulness and credibility judgments of the review set might be reduced when consumers are more tolerant of contradictions and uncertainties, when the attributes of concern are not deemed relevant, or when aggregated ratings are more salient. Future research can explore these other possible boundary conditions. We also acknowledge that cross-review incoherence is only one source of purchase deferral, and there may be other antecedents, again suggesting ample opportunities for future research.

Finally, it is possible that review set credibility and helpfulness have a more complex relationship than what we theorize in the present paper. Helpfulness and credibility are distinct theoretical constructs (Cheung and Thadani 2012), and they are rarely examined simultaneously (e.g., Cheung et al. 2012, Hong et al. 2017). The extent to which one may have a causal impact on the other is an interesting opportunity for future research, particularly given the important roles that these constructs independently play in consumers' decision-making process (Cheung and Thadani 2012).

Conclusion

This paper examines how, why, and when cross-review incoherence influences consumers' purchase deferral. In two laboratory experiments, we find that the presence of cross-review incoherence reduces attitude certainty and increases purchase deferral through lower perceived helpfulness and credibility of the review set and that these pathways are weakened when the specificity of contextual information is high. This work illustrates exciting opportunities for future research on other

antecedents of purchase deferral as well as on the intricate and nuanced interrelationships among reviews.

Acknowledgments

The authors are grateful to the senior editor, the associate editor, and three anonymous reviewers for their constructive guidance and suggestions during the review process. They thank Qianqian Ben Liu for sharing the review stimuli from his (2017) paper coauthored with Elena Karahanna, which we used to construct our review stimuli in the experiments. They are also grateful to K. Doreen MacAulay and Dennis Walpole for their help in recruiting experiment participants.

Endnotes

¹ For simplicity, we refer to “products” instead of “products and services” in our theorizing, but our arguments are applicable to reviews of both products and services.

² An attribute is typically at a granular level with a certain value for a particular product. For example, the opinion that “the laptop has a big screen and a heavy weight” involves two attributes: screen size (big) and weight (heavy).

³ Once participants indicated they were ready to make a decision, we used a three-item scale to measure how likely they were to purchase the product. The descriptive results are provided in Online Appendix D.

⁴ Although attitude about the product was not part of our conceptual model, it was necessary to include a measure of this in order for participants to report their attitude certainty. Attitude toward the product was measured with three items on a semantic differential scale (Schlosser 2011). Participants were asked, “Based on the reviews, what is your overall opinion of camera Model X?” An example item was from “very bad” to “very good.” Cronbach's alpha was $\alpha = 0.91$. Because this variable is not relevant to our hypotheses, we do not discuss it further.

⁵ We also examined the main effects of our manipulated factors on attitude certainty although they are outside the scope of the present paper. The presence of incoherence had a significant negative effect on attitude certainty, $t(735) = -2.14$, $p = 0.033$, Cohen's $d = -0.16$, whereas context specificity had a nonsignificant positive effect, $t(735) = 0.76$, $p = 0.447$, Cohen's $d = 0.06$.

References

- Ahn H-Y, Wu L, Kelly S, Haley E (2011) A qualitative study of college student responses to conflicting messages in advertising: Anti-binge drinking public service announcements versus wine promotion health messages. *Internat. J. Public Health* 56(3):271–279.
- Alba J, Lynch J, Weitz B, Janiszewski C, Lutz R, Sawyer A, Wood S (1997) Interactive home shopping: Consumer, retailer, and manufacturer incentives to participate in electronic marketplaces. *J. Marketing* 61(3):38–53.
- Alter AL, Oppenheimer DM (2009) Uniting the tribes of fluency to form a metacognitive nation. *Personality Soc. Psych. Rev.* 13(3):219–235.
- Aronson E, Wilson TD, Brewer MB (1998) Experimentation in social psychology. Gilbert DT, Fiske ST, Lindzey G, eds. *The Handbook of Social Psychology*, vol. 1–2, 4th ed. (McGraw-Hill, New York), 99–142.
- Babić Rosario A, Sotgiu F, De Valck K, Bijmolt THA (2016) The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors. *J. Marketing Res.* 53(3):297–318.

- Benlian A, Titah R, Hess T (2012) Differential effects of provider recommendations and consumer reviews in e-commerce transactions: An experimental study. *J. Management Inform. Systems* 29(1):237–272.
- Brashier NM, Marsh EJ (2020) Judging truth. *Annual Rev. Psych.* 71(1):499–515.
- Cheung CMK, Thadani DR (2012) The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems* 54(1):461–470.
- Cheung CM-Y, Sia C-L, Kuan KKY (2012) Is this review believable? A study of factors affecting the credibility of online consumer reviews from an ELM perspective. *J. Assoc. Inform. Systems* 13(8):618–635.
- Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.
- Choudhury V, Sampler JL (1997) Information specificity and environmental scanning: An economic perspective. *Management Inform. Systems Quart.* 21(1):25–53.
- Darley JM, Gross PH (1983) A hypothesis-confirming bias in labeling effects. *J. Personality Soc. Psych.* 44(1):20–33.
- DeMotta Y, Chao MC-H, Kramer T (2016) The effect of dialectical thinking on the integration of contradictory information. *J. Consumer Psych.* 26(1):40–52.
- Dhar R (1996) The effect of decision strategy on deciding to defer choice. *J. Behav. Decision Making* 9(4):265–281.
- Dhar R, Nowlis SM (2004) To buy or not to buy: Response mode effects on consumer choice. *J. Marketing Res.* 41(4):423–432.
- Dridi R, Zammali S, Alsulimani T, Arour K (2020) Effective rating prediction based on selective contextual information. *Inform. Sci.* 510:218–242.
- Elliot AJ, Devine PG (1994) On the motivational nature of cognitive dissonance: Dissonance as psychological discomfort. *J. Personality Soc. Psych.* 67(3):382–394.
- Festinger L (1962) *A Theory of Cognitive Dissonance* (Stanford University Press, Stanford, California).
- Fishbein M, Ajzen I (1975) *Intention and Behavior: An Introduction to Theory and Research* (Addison-Wesley, Reading, MA).
- Floyd K, Freling R, Alhoqail S, Cho HY, Freling T (2014) How online product reviews affect retail sales: A meta-analysis. *J. Retailing* 90(2):217–232.
- Goodman JS, Wood RE, Hendrickx M (2004) Feedback specificity, exploration, and learning. *J. Appl. Psych.* 89(2):248–262.
- Greenleaf EA, Lehmann DR (1995) Reasons for substantial delay in consumer decision making. *J. Consumer Res.* 22(2):186–199.
- Haselhuhn MP, Schweitzer ME, Wood AM (2010) How implicit beliefs influence trust recovery. *Psych. Sci.* 21(5):645–648.
- Hayes AF (2015) An index and test of linear moderated mediation. *Multivariate Behav. Res.* 50(1):1–22.
- Hayes AF (2018) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (Guilford Press, New York).
- Hong Y, Pavlou PA (2014) Product fit uncertainty in online markets: Nature, effects, and antecedents. *Inform. Systems Res.* 25(2):328–344.
- Hong H, Xu D, Wang GA, Fan W (2017) Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems* 102:1–11.
- Hu N, Pavlou PA, Zhang J (2017) On self-selection biases in online product reviews. *Management Inform. Systems Quart.* 41(2):449–471.
- Jabr W, Liu B, Yin D, Zhang H (2020) Research curation on online word-of-mouth. Bush A, Rai A, eds. *MIS Quart. Res. Curations*. Accessed November 23, 2020, <https://www.misqresearchcurations.org/blog/2020/11/23/online-word-of-mouth>.
- Jensen ML, Averbach JM, Zhang Z, Wright KB (2013) Credibility of anonymous online product reviews: A language expectancy perspective. *J. Management Inform. Systems* 30(1):293–324.
- Klein KJ, Kozlowski SWJ (2000) From micro to meso: Critical steps in conceptualizing and conducting multilevel research. *Organ. Res. Methods* 3(3):211–236.
- Krijnen JMT, Zeelenberg M, Breugelmans SM (2015) Decision importance as a cue for deferral. *Judgment Decision Making* 10(5):407–415.
- Ledgerwood A, Eastwick PW, Smith LK (2018) Toward an integrative framework for studying human evaluation: Attitudes toward objects and attributes. *Personality Soc. Psych. Rev.* 22(4):378–398.
- Lei Z, Yin D, Zhang H (2021) Focus within or on others: The impact of reviewers' attentional focus on review helpfulness. *Inform. Systems Res.* 32(3):801–819.
- Lei Z, Yin D, Zhang H (2022) Swayed by the reviews: Disentangling the effects of average ratings and individual reviews in online word-of-mouth. *Production Oper. Management*. Forthcoming.
- Liu QB, Karahanna E (2017) The dark side of reviews: The swaying effects of online product reviews on attribute preference construction. *Management Inform. Systems Quart.* 41(2):427–448.
- Liu RR, Zhang W (2010) Informational influence of online customer feedback: An empirical study. *J. Database Marketing Customer Strategy Management* 17(2):120–131.
- Liu X, Lee D, Srinivasan K (2019) Large-scale cross-category analysis of consumer review content on sales conversion leveraging deep learning. *J. Marketing Res.* 56(6):918–943.
- Morrison M (2011) One phenomenon, many models: Inconsistency and complementarity. *Stud. Hist. Philos. Sci.* 42(2):342–351.
- Mudambi SM, Schuff D (2010) What makes a helpful online review? A study of customer reviews on Amazon.com. *Management Inform. Systems Quart.* 34(1):185–200.
- Novemsky N, Dhar R, Schwarz N, Simonson I (2007) Preference fluency in choice. *J. Marketing Res.* 44(3):347–356.
- Pang J, Keh HT, Li X, Maheswaran D (2017) “Every coin has two sides”: The effects of dialectical thinking and attitudinal ambivalence on psychological discomfort and consumer choice. *J. Consumer Psych.* 27(2):218–230.
- Peng C-H, Yin D, Zhang H (2020) More than words in medical question-and-answer sites: A content-context congruence perspective. *Inform. Systems Res.* 31(3):913–928.
- Purnawirawan N, De Pelsmacker P, Dens N (2012) Balance and sequence in online reviews: How perceived usefulness affects attitudes and intentions. *J. Interactive Marketing* 26(4):244–255.
- Qahri-Saremi H, Montazemi AR (2019) Factors affecting the adoption of an electronic word of mouth message: A meta-analysis. *J. Management Inform. Systems* 36(3):969–1001.
- Qiu L, Pang J, Lim KH (2012) Effects of conflicting aggregated rating on eWOM review credibility and diagnosticity: The moderating role of review valence. *Decision Support Systems* 54(1):631–643.
- Reber R, Schwarz N (1999) Effects of perceptual fluency on judgments of truth. *Consciousness Cognition* 8(3):338–342.
- Rucker DD, Tormala ZL, Petty RE, Briñol P (2014) Consumer conviction and commitment: An appraisal-based framework for attitude certainty. *J. Consumer Psych.* 24(1):119–136.
- Schlosser AE (2011) Can including pros and cons increase the helpfulness and persuasiveness of online reviews? The interactive effects of ratings and arguments. *J. Consumer Psych.* 21(3):226–239.
- Schwanenflugel PJ, Kipp Harnishfeger K, Stowe RW (1988) Context availability and lexical decisions for abstract and concrete words. *J. Memory Language* 27(5):499–520.
- Schwarz N, Newman E, Leach W (2016) Making the truth stick & the myths fade: Lessons from cognitive psychology. *Behav. Sci. Policy* 2(1):85–95.
- Shan G, Zhou L, Zhang D (2021) From conflicts and confusion to doubts: Examining review inconsistency for fake review detection. *Decision Support Systems* 144:113513.

- Shiu EMK, Walsh G, Hassan LM, Shaw D (2011) Consumer uncertainty, revisited. *Psych. Marketing* 28(6):584–607.
- Shoham M, Moldovan S, Steinhart Y (2017) Positively useless: Irrelevant negative information enhances positive impressions. *J. Consumer Psych.* 27(2):147–159.
- Solomon LJ, Rothblum ED (1984) Academic procrastination: Frequency and cognitive-behavioral correlates. *J. Counseling Psych.* 31(4):503–509.
- Sun M (2012) How does the variance of product ratings matter? *Management Sci.* 58(4):696–707.
- Tormala ZL, Petty RE (2004) Source credibility and attitude certainty: A metacognitive analysis of resistance to persuasion. *J. Consumer Psych.* 14(4):427–442.
- Tseng S, Fogg BJ (1999) Credibility and computing technology. *Comm. ACM* 42(5):39–44.
- Tversky A, Shafir E (1992) Choice under conflict: The dynamics of deferred decision. *Psych. Sci.* 3(6):358–361.
- Urbany JE, Dickson PR, Wilkie WL (1989) Buyer uncertainty and information search. *J. Consumer Res.* 16(2):208–215.
- Wathen CN, Burkell J (2002) Believe it or not: Factors influencing credibility on the web. *J. Amer. Soc. Inform. Sci. Tech.* 53(2):134–144.
- Yin D, Bond SD, Zhang H (2014) Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *Management Inform. Systems Quart.* 38(2):539–560.
- Yin D, Bond SD, Zhang H (2017) Keep your cool or let it out: Non-linear effects of expressed arousal on perceptions of consumer reviews. *J. Marketing Res.* 54(3):447–463.
- Yin D, Bond SD, Zhang H (2021) Anger in consumer reviews: Unhelpful but persuasive? *Management Inform. Systems Quart.* 45(3):1059–1086.
- Yin D, Mitra S, Zhang H (2016) When do consumers value positive vs. negative reviews? An empirical investigation of confirmation bias in online word of mouth. *Inform. Systems Res.* 27(1):131–144.
- Zhang Z, Yang K, Zhang JZ, Palmatier RW (2022) Uncovering synergy and dysynergy in consumer reviews: A machine learning approach. *Management Sci.* Forthcoming.