City Research Online

City St George's, University of London

Citation: van Laer, T., Ludwig, S. & Escalas, J.E. (2017). Understanding the value of stories in experiential reviews. Cambridge, Massachussets: Marketing Science Institute.

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/18846/

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Understanding the Value of Stories in Experiential Reviews:

Working Paper

TOM VAN LAER, STEPHAN LUDWIG, AND JENNIFER EDSON ESCALAS

INTRODUCTION

Seventy percent of consumers indicate that they base purchase decisions on online consumer reviews, thus marking these reviews as the most influential form of word of mouth (WOM) after recommendations from family and friends (Nielsen, 2015). In online reviews, ordinary consumers (i.e., reviewers) write about purchases, and a web hosting site aggregates these evaluative texts into an organized format (McQuarrie et al., 2015). Most hosting sites offer consumers the option to respond to and evaluate the review. For example, on Yelp (2017), consumers are asked whether reviews are useful, funny, or cool, while on TripAdvisor (2017), each review is followed by the question "Helpful?" and a button to make a thumbs-up gesture. This positive feedback, that is, the attitudinal response to the review, raises a review's ranking and visibility on the site and may change consumers' attitudes and purchase decisions (Moore, 2015). In this paper, we propose and find support for an overlooked feature of reviews that influences consumer behavior: the extent to which reviews tell stories.

This paper offers a narratological perspective on reviews following Jurafsky et al. (2014), who maintain that reviews' content is "overwhelmingly focused on narrating experiences … rather than discussing." A review consists of an account of a sequence of events leading to a transition in a character from an initial state to a later state (i.e., a story, Bennett and Royle,

2004) in which the reviewer is the main character. Thus, we answer the calls for more specific content analyses to examine how stories create value for consumers (Moore, 2012, Levy, 2006, Figueiredo and Scaraboto, 2016).

Prior research (Anderson and Simester, 2014, Holbrook and Grayson, 1986, Russell, 2002, van den Hende et al., 2012, Barnhardt et al., 2016) provides a valid description of the impact of stories featuring material purchases. "Material purchases are those made with the primary intention of acquiring a material good: a tangible object that is kept in one's possession" (Van Boven and Gilovich, 2003, p. 1194), such as a battery, a digital camera, or a watch. In these stories, a material good plays the role of a character or prop that supports the plot. Both product placement (Russell, 2002) and product concept test (van den Hende et al., 2012) research consistently finds strong support for the degree of connection between the material good and the story plot as the determinant of attitudes toward the story and the good.

We propose that prior research does not fully account for the influence of stories featuring experiential purchases on consumer behavior; a gap we intend to fill. Experiential purchases "are those made with the primary intention of acquiring a life experience" (Van Boven and Gilovich, 2003, p. 1194), such as visiting a bar, attending a class, or playing a sport. Unlike material purchases, experiential purchases do not play a supporting role in a story of an experiential purchase. Instead, the entire story describes the recounted consumption experience (Pechmann and Wang, 2010). To investigate the content and consequences of reviews of experiential purchases, we use a multimethod approach. In doing so, we aim to make three contributions to the narratology and WOM language literature streams.

First, we demonstrate a reliable and valid automatic assessment of narrative elements from big (textual) data. Essentially, the various theories in cognitive psychology (Bruner, 1986,

Green and Brock, 2000, Green and Brock, 2002), linguistics (Mandler, 1984, Brewer and Lichtenstein, 1981), social psychology (Gergen and Gergen, 1988), and rhetoric (Burke, 1962) remain at the conceptual level of framework building without extending to the empirical level of stories as bodies of texts that can be measured systematically and quantitatively. Therefore, this paper offers an automated text analysis of TripAdvisor to facilitate a more in-depth, comprehensive understanding of consumer stories. By means of computational linguistics, which is the branch of linguistics related to automated analysis of large collections of digital texts (big data corpora, McEnery and Hardie, 2011), various narrative elements are manifested.

Second, we show that the more reviews of experiential purchases include narrative elements, the more they receive positive feedback. Not all reviewers are equally able or motivated to reconstruct their experience into an engaging story, and this inequality in narrativity (Sturgess, 1992, Porter Abbott, 2002) is likely to contribute to additional variance in positive feedback. As we elaborate subsequently, an experience review is more narrative when its characters are more identifiable, its story plot is more imaginable, and its genre's emotional story shape is more fluctuating. These narrative elements are likely to increase positive feedback.

Third, with two experimental studies we establish that narrative transportation into reviews drives consumer responses, including attitude toward and intention to purchase the reviewed experience. In other words, our research has implications not only for positive feedback but also for consumer behavior in the overall marketplace. Because of the growing popularity and impact of online consumer reviews, it is empirically important to understand their influence on how consumers enter the narrative worlds evoked by reviews (i.e., narrative transportation, Green, 2008) and make consequential decisions.

LITERATURE REVIEW

The emerging literature on WOM language attempts to explain consumer responses to online reviews (Kronrod and Danziger, 2013, Berger and Iyengar, 2013, Berger, 2014, Chen and Berger, 2016, Schellekens et al., 2010, Moore, 2015). These responses can stem from reliance on more superficial, contextual cues or from attending to the review content. If consumers rely on contextual cues, their evaluation tends to be more positive if the purchase ranks higher (Scott and Orlikowski, 2012); reviews are newer and scarcer (Ludwig et al., 2013), less extreme (Cao et al., 2011, Mudambi and Schuff, 2010), literate (Vásquez, 2014), negative (Yin et al., 2014), readable (Ghose and Ipeirotis, 2011), and longer (Pan and Zhang, 2011); and reviewers are more experienced (Godes and Mayzlin, 2004) and disclose their current city, country, and name (or nom de plume, Forman et al., 2008). If consumers attend to the review content, their evaluation depends on the focus of the story: a material or an experiential good. Whereas researchers (Anderson and Simester, 2014, Holbrook and Grayson, 1986, Russell, 2002, van den Hende et al., 2012, Barnhardt et al., 2016) generally accept that plot connection is the crucial determinant of consumer responses to stories of material goods, several narrative elements may jointly drive the responses to experience reviews. A consumption experience is "an event or series of events that one lives through" (Van Boven and Gilovich, 2003, p. 1194). As such, we propose that consumption experiences provide the building blocks for stories: identifiable characters (Slater and Rouner, 2002), imaginable plot (Green and Brock, 2002), and genre (Gergen and Gergen, 1988). Thus, experience reviews are the reconstructions of consumption experiences into stories that consumers-as-reviewers tell other consumers, and their elements are the units that lead to narrative persuasion.

Extant research in cognitive psychology (Bruner, 1986, Green and Brock, 2000, Green and Brock, 2002) and consumer research (Phillips and McQuarrie, 2010) demonstrates that a story can engross consumers in a transformational experience, which is captured and conceptualized in narrative transportation: "the extent to which (1) a consumer empathizes with the story characters and (2) the story plot activates his or her imagination, which leads him or her to experience suspended reality during story reception" (van Laer et al., 2014, pp. 799-800). Narrative transportation leads to narrative persuasion of the consumer (Escalas, 2007, Green and Brock, 2000, Green and Brock, 2002, Slater and Rouner, 2002). Transporting stories are perceived to be similar to real-life experiences (Green, 2004, Escalas, 2004). In the context of experience reviews therefore, narrative transportation provides consumers with the perception that their resulting evaluation is based on direct experience, which typically makes novel information easier to understand and seemingly intuitively truthful (Marsh and Fazio, 2006). Consumers perceive to have gained a more realistic, vicarious "foretaste" of the affective and cognitive consequences that are associated with making the experiential purchase (Woodside et al., 2008).

Thus, narrative transportation leads to a positive attitude toward the review itself. Giving positive feedback on the review (e.g., a helpfulness vote) is the manifestation of consumers' positive attitude toward the review and the experience described therein. The foretaste feature of narrative transportation also facilitates the formation of story-consistent attitudes and behavioral intentions (Argo et al., 2008, Escalas, 2007, Wang and Calder, 2006). Holding story valence and all else constant, a positive attitude toward the review should therefore predict a positive attitude toward the reviewed experience. Furthermore, given that consumers typically show purchase

intentions consistent with their attitudes (Schlosser, 2003), favorable attitudes toward the reviewed experience should extend to favorable purchase intentions as well. Thus:

H1: (a) Narrative transportation predicts positive feedback, (b) which predicts attitude toward the reviewed experience, (c) which in turn predicts purchase intention.

To understand how an experience review can transport and further affect consumers, we conceptually break down its narrative building blocks—identifiable characters, imaginable plot, and genre. We highlight two elements that can increase characters' identifiability (landscape of affective consciousness and landscape of cognitive consciousness), three elements that can increase plot imaginability (temporal embedding, spatial embedding, and drama), and five story shapes that can change genre's emotionality (progressive, regressive, stable, comedy, and tragedy), and we provide examples using sentences from illustrative TripAdvisor reviews of "things to do" (i.e., experiences) in Las Vegas.

Landscapes of Affective and Cognitive Consciousness

A story is an account of a sequence of events leading to a transition in a character from an initial state to a later state (Bennett and Royle, 2004) and thus, by its very essence, gives some insight into a character's feelings and thoughts (Bruner, 1986). Following Bruner (1986) and Feldman et al. (2014), we define the landscape of affective consciousness as the extent to which an experience review recounts initial events about which characters express feelings that, in turn, lead to subsequent events. For example, a review of *Kà*, a circus show in Las Vegas, includes "There was a lot of action. That I love in this show. I would totally go see it again." Similarly,

we define the landscape of cognitive consciousness as the extent to which an experience review recounts initial events about which a character expresses thoughts that, in turn, lead to subsequent events. For example, a review of *Vegas! The Show*, a musical, includes: "They changed the show!!! I think the 'old' show was more complete. If they don't bring back the original show, this is my last time attending this show!" Feldman et al. (2014) suggest that the landscapes of affective and cognitive consciousness make a narrative more transporting, affecting consumer responses: When stories have well-developed landscapes of consciousness, consumers make more inferences and exert more effort to identify with the characters. Thus:

H2: The more an experience review gives insight into (a) what a character is feeling and (b) what a character is thinking, the more it affects consumer responses.

Temporal Embedding

A story plot's temporal embedding is the result of (1) narrative movement—that is, the temporal flow of the events indicating the direction of the story (Woermann and Rokka, 2015)—and (2) narrative framing—that is, the thematic and symbolic parallels among different events in the story (Thompson, 1997). First, narrative movement organizes events in terms of a temporal dimension: Things occur over time with some sort of beginning, middle, and end. Second, narrative framing establishes a network of relationships between story characters and events that allows for making causal inferences (Escalas, 1998).

Scholars from various fields, including consumer research, have debated what constitutes narrative movement and framing (Adaval and Wyer, 1998, Adaval et al., 2007, Barthes, 1975, Green and Brock, 2002). They contend that past–present–future causal chains are necessary to

translate texts into stories with imaginable plots that transport and affect consumer responses. For example, another *Vegas! The Show* review includes "The first half seemed to drag on until the bird trainer and his buddies came on. Because they were hilarious and their performance seemed to add life to the show and energize the crowd. The second half of the show was a lot of fun!!"

Spatial Embedding

Spatial embedding is the extent to which the text focuses on particular events, rather than developing abstractions, categorizations, and/or generalizations. Spatial embedding has a more narrowly defined format and function than schemata or scripts. A schema reflects the universal knowledge of a particular domain (Alba and Hasher, 1983), whereas scripts are mental representations of common events as abstractions (Abelson, 1981). Consumers give low narrativity ratings to texts conforming strictly to the latter (Brewer and Lichtenstein, 1981), and reviews' valence rather than content determines their purchase intention (Schellekens et al., 2010). Stories with an imaginable plot that affects consumer responses may not need to offer a camera-recorded view of space, as some movies and video games do, but they are also not as abstract as schemata or scripts. For example, a *Titanic: The Artifact Exhibition* review includes "They have lots of plates from the ship, replicas of the 'bedrooms' for the 3rd class and 1st class passengers. They have a real (freshwater) iceburg [*sic*] and a large section of the boat."

Drama

Stories' drama emerges from oddities or twists that breach Burke's (1962) "dramatistic pentad." He asserts that any complete story should answer the following five questions, which correspond to the five elements of the pentad: What was done? (act), Who did it? (agent), How did he or she do it? (agency), Where and when was it done? (scene), and Why? (purpose). The pentad is breached, for example, if actions are not on purpose or agents and agency do not match. As another *Kà* review reads "I was sometimes confused on who the characters were and the plot was sometimes strange." Berger and Iyengar (2013) show that the written modality of online reviews fosters telling of such interesting content. If consumers strive to understand the breaches, they may mentally engage in a plot that transports and affects their responses (Nielsen and Escalas, 2010).

In summary, a story may evoke more narrative transportation, and consequently affect consumer responses, the more its plot is temporally and spatially embedded and dramatic. Thus:

H3: The more an experience review's story plot shows (a) temporality and (b) spatiality of events, and (c) breaches in the dramatistic pentad, the more it affects consumer responses.

Genre

In line with Genette (1979/1992), we define genre as the different story shapes that emerge from culturally determined conventions in a given society at a given time. Gergen and Gergen (1988) classify genres in a taxonomy of five basic types: progressive, regressive, stable, comedy, and tragedy. In a progressive genre, events continuously improve for characters over the course of the story line, while in a regressive genre, events decline over the course of the story

line. In a stable genre, events neither improve nor decline significantly over the course of the story line. The final two genres involve emotional slopes that alternate in sign—that is, story shapes that rise and decline (or decline and then rise) over the course of the story line (Vonnegut, 2005). In the case of a comedy, events start out favorable, deteriorate, and then end on a positive note. Thus, it is a regressive slope, followed by a progressive slope. An example is the following review of *Mystery Adventures*, a live action role-playing game organized in Las Vegas:

> This is definitely an unusual thing to do in Las Vegas, but can be a wonderful change of pace. If you are into CSI and like solving mysteries, this is for you. If you'd rather just kick back, this might be a bit much. Max seemed nervous at first with lots of 'uhhh's and ummmms, but warmed up quickly. The mystery started out slow ... which might be natural, but picked up pace and excitement as the night went on. And it did go on ... from 7pm to well past 10pm. Very exciting and worth the effort we put into it.

The opposite of this form is a tragedy, which is a progressive slope followed by a regressive slope. When the genre is a tragedy, characters have almost attained their goal and then are brought low. An example is another *Mystery Adventures* review:

> After attending, I was disappointed. First of all you have to travel off the Strip to get to the location. It would be much more convenient if they came and picked you up. I was expecting an exciting adventure but found Mystery Adventures to be dull. The first crime scene was the best. It was thought stimulating. After that, it went down hill [*sic*].

Emotional story shapes that change over the course of a story line are more engaging than those that do not alternate in sign (Vonnegut, 2005). Because of comedy's downs and ups and tragedy's ups and downs, it is reasonable to expect consumers to experience narrative transportation and respond accordingly. Following Gergen and Gergen (1988), we therefore

develop a precise hypothesis, to anticipate the effects of different genres on narrative transportation and consequential consumer responses:

H4: An experience review's genre that involves a changing emotional story shape (i.e., a comedy or tragedy) affects consumer responses more than a genre that involves a progressive, regressive, or stable shape.

We address these four hypotheses in a series of three studies. First, study 1 consists of an automated text analysis of TripAdvisor reviews to test hypotheses 2a–4. Second, study 2, which tests hypotheses 1a and 2a–4, had participants evaluate systematic subsample of these reviews. Third, study 3 is a controlled experiment that addresses hypotheses 1a–1c and 3c with reviews that we handcrafted to further clarify and extend the findings of studies 1 and 2.

STUDY 1

Method

Sampling Frame and Text-Mining Procedure. We derived a corpus of experience reviews using the KoNstanz Information MinEr software program (Berthold et al., 2008), which we ran to access and parse the publicly available HTML (Hypertext Markup Language) and XML (Extensible Markup Language) pages on http://www.tripadvisor.com. Our sample included all English reviews of "things to do" (i.e., experiences) in Las Vegas, which has the world's most purchased experiences with 39,668,221 annual visitors (Las Vegas Convention and Visitors Authority, 2017, Love Home Swap, 2015), posted between TripAdvisor's founding in February

2000 and our text mining. We excluded reviews in languages other than English because of the difficulties of interlingual comparison in automated text analysis.

The corpus stemmed from TripAdvisor for several reasons. First, TripAdvisor is the most inclusive, dedicated hosting site for reviews of experiences, such as visiting a restaurant, attending a performance, playing a game, or taking a sightseeing tour (Scott and Orlikowski, 2012, p. 29). Second, on TripAdvisor any consumer can give positive feedback on a review by voting it helpful with a thumbs-up gesture. Consumers cannot give negative feedback. Third, on TripAdvisor reviewers post reviews of leisure travel–related experiential purchases and, in doing so, make sense of their experiences by narrating the events they go through on social media (van Laer and de Ruyter, 2010). When people establish relationships between characters and events, stories are created (Somers, 1994). Reviews on TripAdvisor are therefore personal stories of experiences (Bennett and Royle, 2004). Fourth, the website allows us to control for eight contextual cues that might affect positive feedback:

1. Experience rank order: TripAdvisor's Popularity Index calculates and presents this ranking, which is determined using a proprietary algorithm that incorporates reviews' numerical rating (TripAdvisor Support, 2017). The rating ranges from 1 ("terrible") to 5 ("excellent"). We measured experience rank order as the experience's order number divided by the total number of experiences in the category. As reviews of the same experience may also be correlated, we also included dummy variables for experiences. These dummies also absorb the within-experience rank order variation.
2. Pictures: the number of visitor photos per experience. Adaval and colleagues (2007, 1998) show that the presence of pictures affects the processing of experiential information.

3. Review age: the number of days the review has been online. Older reviews have a greater opportunity to receive positive feedback (Ludwig et al., 2013).
4. Review eloquence: the proportion of long words (>6 letters). Long words increase perceived eloquence, which increases positive feedback (Vásquez, 2014).
5. Review extremity: reviews' rating ranges from 1 ("terrible") to 5 ("excellent"). Cao et al. (2011) indicate that reviews with extreme ratings receive more positive feedback. We measured review extremity as the difference between the review's rating and the reviewed experience's average rating.
6. Review readability: the proportion of nonfluencies (e.g., er, hm, umm). Nonfluencies reduce readability, which decreases positive feedback (Ghose and Ipeirotis, 2011).
7. Review volume: the number of reviews per experience. A higher number of reviews per experience decreases a review's competitiveness for positive feedback (Ludwig et al., 2013).
8. Reviewer expertise: measured as the total number of helpfulness votes a reviewer received. Because expertise may be a necessary condition for writing transporting reviews (Godes and Mayzlin, 2004), a higher helpfulness vote's volume may be associated with more positive feedback.

Noting the potential influence of hidden reviewer identity (Ghose and Ipeirotis, 2011, Forman et al., 2008), we also mined whether reviewers' current location and name or nom de plume were displayed. However, almost all reviewers reported these (90.76% and 99.45%, respectively); thus, these variables lacked sufficient variance to be of informational value, so we excluded them from the analysis. We accounted for varying review lengths (Pan and Zhang, 2011) in our operationalization of the narrative elements.

Our procedure mined 190,461 reviews of 989 experiences consumed in Las Vegas. The reviews averaged 89 words (SD = 89.74; ranging from 3 to 2,335 words), seven sentences (SD = 4.56; ranging from 1 to 148 sentences), and .77 positive feedback (SD = 2.01; ranging from 0 to 103 votes) and included 65.25% reviews without votes. We discuss the lack of positive feedback in the section on limitations.

Narrative Elements Operationalization. We conducted an automated text analysis of *n*-grams of multiple word lengths. A set of *n*-grams in a text is the set of all distinct sequences of *n* words (Vásquez, 2014). In support of our analysis, we relied on the Linguistic Inquiry and Word Count (LIWC) software program. LIWC measures the intensities with which word categories, classified in dictionary entries, are used in each text. Since Pennebaker et al.'s (2007) quantitative operationalization, more than 120 studies have employed the software (Tausczik and Pennebaker, 2010). We used validated LIWC dictionary entries, which Pennebaker et al. developed, as a starting point from which to operationalize our narrative elements, except for drama for which we developed a dictionary from scratch. We provide definitions, operationalizations, and representative articles of the narrative elements in table 1 (Tables and Figures follow "References" throughout).

Linguistically, landscapes of consciousness consist of three consecutive parts: motion (an initiating event), a psychological state (affective or cognitive), and another motion (a sequential event). Thus, to measure the quality of landscapes of affective and cognitive consciousness, we broke each review into sentences and had LIWC count the words in each sentence that matched its "motion" (168 words; e.g., arrive, car, go), "affective processes" (i.e., affective consciousness; 915 words; e.g., abandon, cried, happy), and "insight" (i.e., cognitive consciousness; 195 words; e.g., consider, know, think) dictionary entries. LIWC provides counts

as an intensity: a fraction of the number of words in the designated text. To account for varying review lengths (Pan and Zhang, 2011), we converted these intensities into ratios that signified the number of motion–affective process–motion (landscape of affective consciousness) and motion–insight–motion (landscape of cognitive consciousness) trigrams (word sequences of length three) divided by the number of sentences in a review, respectively.

Time (narrative movement) and causation (narrative framing) strung linguistically together indicate the existence of temporal embedding. Thus, to measure the quality of temporal embedding, we had LIWC count the words in each review that matched its "time" (239 words; e.g., end, season, until) and "causation" (108 words; e.g., because, effect, hence) dictionary entries. We then converted the intensities into a continuous variable that signified the extent to which the review was temporally embedded, coded as (0) neither time nor causation unigrams (word sequences of length one), (1) time unigrams but no causation unigrams, (2) causation unigrams but no time unigrams, or (3) time and causation unigrams in a review. As causation words imply chronology but time words do not imply causation, the presence of solely causation unigrams indicated greater temporal embedding than the presence of solely time unigrams.

Linguistically, references to space (a particular event) indicate the quality of spatial embedding. Thus, to measure spatial embedding, we had LIWC count the words in each review that matched its "space" dictionary entry (220 words; e.g., down, in, thin). The intensity represented the proportion of space unigrams in a review.

A breach in the dramatistic pentad indicates the existence of drama linguistically. Thus, to operationalize drama, a team of professional lexicographers followed Pennebaker et al.'s (2015) procedure to compile a dictionary related to surprise (32 words; e.g., amaze, astonish, shock, startle, stupefy; see appendix A), backed by Merriam-Webster (2015) and Oxford

Dictionaries (2015). We had LIWC count the words in each review that matched the custom dictionary entry. The intensity represented the surprise unigram in a review.

To reveal the set of five core trajectories that form the building blocks of genre, we used sentiment analysis to map the emotional shape of each review and then used growth-rate modeling to reveal the set of five core trajectories that form the building blocks of genre. The idea behind sentiment analysis is that words have a positive or negative emotional meaning. Words can thus be a measure of the emotional valence of a sentence and how it changes from sentence to sentence. So, measuring the shape of a story is a question of assessing the emotional polarity of a story at each sentence and how it changes. We conducted this analysis on each review. Finally, we used growth-rate modeling and the method of least squares (Jokisaari and Nurmi, 2009) to tease apart the different emotional shapes present in these stories. Appendix B describes all these methods in more technical detail. These techniques indicate the existence of basic emotional story shapes that form the five genres: progressive (n = 793), regressive (n = 4,601), stable (n = 163,576), comedy (n = 17,279), and tragedy (n = 4,212). The reviews we cite in the conceptual elaboration of genre typify the latter two genres (see figure 1).

Positive Feedback Estimation. We report the means, standard deviations, and intercorrelations of the positive feedback, narrative elements, and control variables in table 2. The low magnitude of the intercorrelations (all $\rho s < .70$) indicates the absence of multicollinearity in our analysis. We included the positive feedback as a variable measured by consumers' thumbs-up gestures. To test the potential effects of the narrative elements on positive feedback, we calculated the standardized regression weights β, Wald's chi-square change, and McFadden's pseudo-R-square. To determine the effect sizes, we used the incidence rate ratio

(IRR), or the factor by which the positive feedback would be expected to change if a narrative element was to increase by one standard deviation, *ceteris paribus*.

To account for the excess number of zero positive feedback (Vuong's $Z = 41.45$, $p < .001$) and the dependent variable's skewed distribution (skewness = 10.29; Shapiro–Wilk's $W = .57$, $p < .001$), we conducted a zero-inflated Poisson regression (Dzhogleva and Lamberton, 2014, Greene, 2011), in which the narrative elements explained the positive feedback. Zero-inflated Poisson regressions estimate the binary and the count model jointly. The binary model, which is logit or probit distributed and determines whether a vote is cast, complements the Poisson-distributed count model that predicts the number of votes. The zero-inflated Poisson regression predicted 65.94% (vs. the observed 65.25% reviews without votes). Appendix C describes the Vuong (1989) test and the zero-inflated Poisson regression in more detail.

We also explicitly modeled the two-step process (see appendix D). Specifically, we estimated two models for predicting positive feedback. First, using a logistic regression (binary model), we determined whether a vote was cast. Second, implementing a truncated-at-zero negative binomial regression (count model), we estimated the number of votes, given that at least one was made. Because no hypothesized effects changed consistently throughout both models and the zero-inflated Poisson regression explains more variance, we report the latter in the "Results" section.

Results

Hypotheses Tests. We report the effects of the narrative elements in table 3. The first model consists of the control variables, which explain 14.20% of the variance in positive

feedback (Wald's $\chi^2_{(24)} = 20759.15$, $p < .001$). In the second, third, and fourth models, we entered the elements for identifiable characters (model 2: Wald's $\chi^2_{\text{Change (2)}} = 2630.32$, $p < .001$), imaginable plot (model 3: Wald's $\chi^2_{\text{Change (3)}} = 3294.72$, $p < .001$), and genre (model 4: Wald's $\chi^2_{\text{Change (4)}} = 97.07$, $p < .001$). These explain additional significant proportions of variance in positive feedback (14.73%, 15.34%, and 15.42%, respectively).

Next, we report the control variables and narrative elements, as summarized in model 4. Regarding the effects of the control variables on positive feedback, the effects of review age (β = .35, SE = .02, $p < .001$, IRR = 1.42), eloquence (β = .03, SE = .01, $p < .001$, IRR = 1.03), extremity (β = .16, SE = .02, $p < .001$, IRR = 1.17),), readability (β = –.03, SE = .01, $p < .001$, IRR = .97), and volume (β = –.51, SE = .09, $p < .001$, IRR = .60), and reviewer expertise (β = .06, SE = .00, $p < .001$, IRR = 1.06) are significant. The effects of experience rank order (β = –.06, SE = .08, $p = .446$) and pictures (β = .02, SE = .02, $p = .858$) are not significant.

For the narrative elements, we find that, consistent with hypotheses 2a–3b, more landscape of affective consciousness (β = .07, SE = .01, $p < .001$, IRR = 1.07), landscape of cognitive consciousness (β = .02, SE = .01, $p < .01$, IRR = 1.02), temporal embedding (β = .20, SE = .01, $p < .001$, IRR = 1.22), and spatial embedding (β = .04, SE = .01, $p < .001$, IRR = 1.05) significantly increase positive feedback. Furthermore, reviews coded as comedies (β = .07, SE = .03, $p < .001$, IRR = 1.07) or tragedies (β = .10, SE = .04, $p < .01$, IRR = 1.10) receive more positive feedback than reviews of other genres; thus, we find support for hypothesis 4. However, no significant effect emerges for drama (β = –.02, SE = .01, $p = .165$); thus, we do not find support for hypothesis 3c.

Robustness Checks. Our operationalization cannot easily capture comedies and tragedies in reviews with less than three emotion words or less than three sentences. If many reviews are

of these shapes and sizes however, our model may overestimate their impact. A sensitivity analysis showed that model 4 explained more variance than a reduced model estimated when solely including reviews with more than two emotion words and sentences (McFadden's pseudo-R^2 = .150; see appendix E). We also checked for the predictive performance of model 4. Before initiating classification, we performed a logistic regression analysis on a training corpus—that is, a randomly drawn subsample of reviews used for training the classification algorithm. With model 4's narrative elements and control variables, the logistic regression model predicted whether a review received positive feedback. We then turned these predictions into the algorithm and applied them out-of-sample. We used a large training corpus (n = 171,435) and a smaller holdout corpus (n = 19,026). As training corpora in computational linguistics increase in size, false negatives decrease while the percentage of false positives is not affected (Das and Chen, 2007). Of the 6,667 reviews with and 12,359 reviews without positive feedback, the algorithm predicted the correct classification of 70.2% of the reviews, which is in line with previous text analyses (ranging from 60% to 70%, Das and Chen, 2007). In summary, the model performs well.

Validation Coding. To ensure the validity of the text analysis, two expert coders blind to the hypotheses followed Weber's (2005) procedure to classify the text of a proportional stratified sample of 90 seven-sentence reviews. The coders used an instrument similar to Escalas and Bettman's (2000) and Woodside et al.'s (2008). The instrument instructs coders to read each review carefully and decide which of the described narrative elements are represented. After practicing on 10 example reviews, the coders classified the sampled reviews, and we compared their classification with the automated text analysis. The coders returned only 3.9% false positives (type I errors) and 8.9% false negatives (type II errors). Between the coders and the

classification algorithm, generally acceptable to near-perfect agreement levels (Krippendorff, 2013) were achieved (.67 < Krippendorff's α < .89). The agreement on tragedy was an exception (Krippendorff's α = .49).

In study 1, we used computational linguistics to map narrativity. Expert coders justified and validated this unique approach carefully. Although it seems more difficult for people than for a computer to assess the emotional polarity of a story at each sentence and determine how it shapes a tragedy in comparison with the more common comedy, our validation was successful. Across 190,461 experience reviews, our text analysis shows that narrativity exerts a distinctive effect on positive feedback.

We did not test a specific hypothesis for narrative transportation in study 1. As previously mentioned, we adopt narrative transportation theory to ground our conceptual framework of narrativity. However, from a purely methodological perspective, the inclusion of narrative transportation in study 1 was impossible, as this variable is not available in online corpora of reviews, which precludes measurement. Instead, we develop study 2 to demonstrate that narrative transportation underlies the effect of narrativity on positive feedback.

There were more reviews without positive feedback (65.25%) than would be predicted from the range of positive feedback counts (1–103) in study 1. One explanation could be that the sheer number of reviews of some experiences makes it more difficult for some reviews to be read and to compete for votes. Indeed, the binary model shows that review volume per experience is the strongest predictor of whether a vote is cast (β = –.70, SE = .07, p < .001; see appendix D), and an additional test shows that reviews without votes compete with a significantly higher review volume (M = 5353, SD = 8419.85) than reviews with votes (M = 1891, SD = 3367.17;

$t_{(152628)} = -43.02$, $p < .001$). Study 2 tests whether the narrativity of the reviews with and without votes is also markedly different.

STUDY 2

Method

Participants. Participants received minimum wage for rating online consumer reviews systematically drawn from study 1's big data set; confidentiality was assured. Participants were 304 Amazon Mechanical Turk (MTurk) workers (46.1% female) and were primarily native English speakers (99.0%). The age of the participants ranged from 18 to 77 years, with an average of 33.15 years (SD = 10.10). Most had received a high school diploma (39.5%), 23.0% had earned an associate or vocational degree, and 29.6% and 7.2% had graduated from university with a bachelor's or master's degree, respectively.

Materials and Procedure. After the introduction to the study, the participants saw 10 reviews randomly drawn from the subset of 90 reviews sampled for validation coding in study 1. Ten is the number of reviews most consumers feel they need to read before they can make an informed decision (Eliot and Anderson, 2015). After reading each review, participants responded to narrative transportation and positive feedback measures. After the study, participants answered demographic measures and then were thanked and dismissed.

Measures. We assessed positive feedback with the question from TripAdvisor adapted to a more nuanced measure: "To what extent was this review helpful?" The 7-point Likert-type scale ranged from "not at all" (6.6%) to "very much" (15.3%). The measure of narrative

transportation was based on Green and Brock's (2000) scale. We measured 13 items, such as "While I was reading the review, I could easily picture the events in it taking place" and "After finishing the review, I found it easy to put it out of my mind" (reverse coded; $\alpha = .79$). This 7-point Likert scale ranged from "strongly disagree" to "strongly agree."

To ensure that participants' narrative transportation is a separate construct from our narrativity measure, we used the study 1 narrative elements' scores in our analyses. Table 4 lists the means, standard deviations, and intercorrelations of the key variables.

Results

Homogeneity Check. Our multivariate analysis of variance (MANOVA) for homogeneity indicates that both reviews with and without votes on TripAdvisor are equally capable of evoking narrative transportation and positive feedback (Wilks's $\lambda = .98$, $F(2, 86) = .80$, $p = .451$). Because narrativity exists in reviews without votes on TripAdvisor, another variable, such as review volume, must be inhibiting positive feedback for these reviews.

Hypotheses Tests. We conducted a mixed-effects regression analysis predicting positive feedback from narrative transportation, adding random effects to eliminate any effect of participant demographics. We find that an increase in narrative transportation correlates with an increase in positive feedback ($\beta = 1.26$, SE = .03, $p < .001$), in support of hypothesis 1a.

Next, we conducted mixed-effects regression analyses predicting narrative transportation and positive feedback from the narrative elements, again introducing random effects for the participants. We find that an increase in the following narrative elements significantly increases both narrative transportation (NT) and positive feedback (PF): landscape of affective

consciousness (NT: β = .03, SE = .01, $p < .05$; PF: β = .09, SE = .03, $p < .01$), landscape of cognitive consciousness (NT: β = .05, SE = .01, $p < .001$; PF: β = .09, SE = .02, $p < .001$), temporal embedding (NT: β = .07, SE = .01, $p < .001$; PF: β = .16, SE = .03, $p < .001$), spatial embedding (NT: β = .16, SE = .01, $p < .001$; PF: β = .33, SE = .03, $p < .001$), comedies (NT: β = .09, SE = .01, $p < .001$; PF: β = .18, SE = .03, $p < .001$), and tragedies (NT: β = .03, SE = .01, $p < .05$; PF: β = .13, SE = .03, $p < .001$). In sum, we find further support for hypotheses 2a–3b and hypothesis 4. The support for hypothesis 3c is mixed: Drama has a significant, positive effect on positive feedback (PF: β = .08, SE = .03, $p < .05$) but not on narrative transportation (NT: β = .02, SE = .02, $p = .204$).

In addition, we bootstrapped the indirect effects for the significant, direct effects of the narrative elements on positive feedback per Hayes's (2013b) approach. The bootstrap estimates presented here and in table 4 are based on 500 bootstrap samples. Narrative transportation mediates the relationship between the narrative elements and positive feedback ($.04 \leq$ point estimates $\geq .20$, lower limits of the 95% CIs $\geq .01$, upper limits of the 95% CIs $\leq .23$), except for drama (point estimate = .02, 95% CI = –.01, .06).

Discussion

In study 2, we empirically measure narrative transportation, which predicts positive feedback and helps explain the effect of narrativity. Here, MTurk workers reported narrative transportation and positive feedback for a proportional stratified sample of reviews. Half the sampled reviews were without votes on TripAdvisor, which allowed us to test the prevalence of the effect of narrativity. Reviews without TripAdvisor votes were as transporting and positively

evaluated as reviews with votes. Few counts of unhelpful reviews were present (6.6%). Thus, it appears that other variables, such as high review volume, increased the number of zeros in study 1 beyond prediction and Poisson distribution.

In study 1, drama does not affect positive feedback. In study 2, drama significantly increases positive feedback but has no effect on narrative transportation. It may be that the effect of drama on consumer responses is subtler than our original prediction. We hypothesized that drama would emerge from breaches in Burke's (1962) dramatistic pentad, guided by surprise. Brewer and Lichtenstein's (1982) structural-affect theory provides a possible explanation for our mixed findings. They argue that when crafting a story, storytellers make choices about how to order the story events. Brewer and Lichtenstein distinguish between different event orders, depending on the story's guiding emotion, either surprise or curiosity. Although both surprise and curiosity orders are consonant with Burke's (1962) view that drama emerges from breaches in the dramatistic pentad, in a surprise order, the order of events adds little to the drama, as Brewer and Lichtenstein (1982) exemplify. In contrast, a curiosity order stimulates consumers to exert effort to understand how the opening event came to pass (Brewer and Lichtenstein, 1982). Perhaps a curiosity order is more widely accepted for online reviews than a surprise order because the former is more mentally stimulating. As Hennig-Thurau et al. (2004) document, online reviewers are relatively more likely to order information to stimulate and help consumers. In addition, Kronrod and Danziger (2013) and Moore (2015) explain that reviewers can intuit which order of events is stimulating drama more and that consumers' preferences mirror reviewers' intuitions. These preferences could lead some reviewers to disfavor a surprise order, and therefore they explain the mixed effects of drama on consumer responses in studies 1 and 2. Study 3 assesses the distinction between surprise- and curiosity-order dramas empirically.

Study 3 also addresses the question of how narrativity's impact on narrative transportation into experience reviews influences consumer responses that are important to the marketplace. Specifically, we keep story valence and other things equal, except for narrative transportation, which we manipulate to establish its exclusive relationship to attitude toward and intention to purchase reviewed experiences, in addition to positive feedback.

STUDY 3

Method

Study 3 was an online experiment with a randomized 2 (instruction: narrative or age-10 reading) × 2 (drama: curiosity or surprise order) full-factorial design. Participants read a review and responded to questions about a trip to Agra, India.

Participants. Ninety-one bachelor's and 65 master's in business students (67.3% female) at a large European university participated to partially fulfill a course requirement. Confidentiality was assured. The age of the participants ranged from 18 to 29 years, with an average of 21.29 years (SD = 1.93).

Materials and Procedure. Participants were introduced to the experiment with the preamble to Adaval and Wyer's (1998) travel brochures study, adapted to the digital age: The South Asian Association for Regional Cooperation (an existing organization) wished to determine what people thought about the information that is shared digitally about things to do in Southeast Asia. Several reviews had ostensibly been given to the marketing group of the

university's business school for testing. Participants were told that they would be reading one of these reviews.

Following the introduction, participants were given one of two of Green and Brock's (2000) written instruction sets, referred to as narrative and age-10 reading instructions. Narrative reading instructions, telling participants to simply pay attention, served as the baseline narrative transportation condition. Age-10 reading instructions were intended to undermine narrative transportation; they asked participants to focus on identifying words that a person reading at the age-10 level would not understand. This task does not distract from the content of a story, but it should reduce narrative transportation (Green and Brock, 2000).

Afterward, participants read a review based on Adaval and Wyer's (1998) travel brochure stimulus with the drama following a curiosity or a surprise order (see appendix F). The story describes predominantly desirable features of a trip to Agra, India; however, for believability purposes, one relatively undesirable aspect of the trip is described as well. In the curiosity order, the story first flashes forward, revealing the end: "I did not get any sleep in Agra, home of the Taj Mahal." From that moment, the story flashes back, and the events are described in chronological order ("My holidays started out fine. After I visited the capital of India, Delhi, I moved on to see the Taj Mahal in Agra…."), finishing with the revelation of the cause for the lack of sleep: "It turns out that Agra accommodations are not luxurious and I spent my nights awake on a straw mat." In the surprise order, before the revelation of the cause of the event, the climax occurs: "Up until that moment, my holidays had been fine, but I did not get any sleep in Agra, home of the Taj Mahal." The event necessary to determine the causal chain is only mentioned in the next sentence. The curiosity (surprise) order review counted 126 (131) words and nine (eight) sentences.

After reading the review, participants responded to narrative transportation, positive feedback, attitude toward the reviewed experience, purchase intention, and ancillary and control measures in random order; attention and manipulation checks; and demographic measures. Participants were then thanked and dismissed.

Dependent Measures. The narrative transportation measure (α = .80) was the same as in study 2. To measure positive feedback, participants rated how helpful, useful, and informative (α = .82) they found the review. The three-item 7-point Likert-type scale by Moore (2015) ranged from "not at all" to "very much."

To measure attitude toward the reviewed experience, we used four 7-point semantic differential–type scales (α = .71) by Osgood, Suci, and Tannenbaum (1957) ranging from "bad" to "good," "worthless" to "valuable," "unpleasant" to "pleasant," and "dirty" to "clean." To measure purchase intention, participants estimated the chance (Juster, 1966), likelihood and intention (Moore, 2015), and want (Adaval and Wyer, 1998) to travel to Agra (α = .89). They reported the four estimates on an 11-point Likert-type scale. Juster's item ranged from "no chance, almost no chance [1 in 100]" to "certain, practically certain [99 in 100]." The other items ranged from "not at all" to "very much." Table 5 lists the means, standard deviations, and intercorrelations of the dependent variables across instruction and drama conditions.

Ancillary Measures. In addition to the measures taken to test hypotheses 1a–1c and 3c, we measured piecemeal processing to test an alternative explanation for the consumer responses. We assessed piecemeal processing by using Adaval and Wyer's (1998) one 7-point semantic differential–type scale and two 7-point Likert-type items. The semantic differential–type scale ranged from "You considered the individual aspects of the trip independently of one another and imagined how it would feel to be there" to "You imagined the overall sequence of events that

occurred on the trip rather than thinking about individual aspects of it" (reverse coded). The Likert-type items asked the extent to which participants compared the specific things they wanted to experience in Agra and the extent to which they formed an overall impression of the trip (reverse coded). These items ranged from "not at all" to "very much."

Attention and Manipulation Checks. To check whether participants had read the entire review carefully, they completed four open-ended questions, designed to test recall of information from the review. It included requests for the name of the capital of India and the location of the Taj Mahal.

We used Green and Brock's (2000) instruction manipulation check. The instruction manipulation check contained two items: "I read the review carefully, just like I would read a story or article for fun" and "While reading the review, I was looking for words and sentences that might not be understood by a 10-year-old reader." The 7-point Likert scales ranged from "strongly disagree" to "strongly agree."

We checked the drama manipulation using Bargh and Chartrand's (2000) procedure in which participants answered open-ended questions, starting with general questions ("When you were reading the review, did you notice anything unusual about the text?" and "What did you notice?") and ending with more specific questions ("Did you notice any particular pattern to the sentences that were included in the review?" and "What particular pattern did you notice?").

Results

Attention and Manipulation Checks. We dropped two participants who gave wrong answers to all four questions on the recall measure from all analyses because they are likely to

have read the review partially or carelessly. Two independent-samples *t*-tests revealed significant differences in the expected direction for the instruction manipulation checks. Narrative condition participants (M = 5.51, SD = 1.62) read the review like a story or article for fun significantly more than age-10 participants (M = 4.69, SD = 1.87; $t_{(152)} = 2.93$, $p < .01$). In turn, age-10 condition participants (M = 6.09, SD = 1.62) looked for words and sentences that might not be understood by a 10-year-old reader significantly more than narrative condition participants (M = 2.79, SD = 1.93; $t_{(150)} = 11.55$, $p < .001$). No participant indicated awareness of the review's drama manipulation. In summary, both manipulations were successful.

Hypotheses Tests. We analyzed narrative transportation (NT), positive feedback (PF), attitude toward the reviewed experience (A_{Exp}), and purchase intention (PI) with a 2 × 2 MANOVA with instruction and drama as between-subjects factors. The results revealed a main effect of instruction at both the multivariate level (Wilks's $\lambda = .78$, $F(4, 147) = 10.38$, $p < .001$) and most univariate levels (NT: $F(1, 150) = 12.67$, $p < .001$; PF: $F(1, 150) = 22.20$, $p < .001$; A_{Exp}: $F(1, 150) = .05$, $p = .821$; PI: $F(1, 150) = 9.27$, $p < .01$) and a main effect of drama at both the multivariate level (Wilks's $\lambda = .63$, $F(4, 147) = 22.01$, $p < .001$) and univariate levels (NT: $F(1, 150) = 40.94$, $p < .001$; PF: $F(1, 150) = 24.65$, $p < .001$; A_{Exp}: $F(1, 150) = 41.60$, $p < .001$; PI: $F(1, 150) = 27.74$, $p < .001$). An interaction between instruction and drama at both the multivariate level (Wilks's $\lambda = .85$, $F(4, 147) = 6.68$, $p < .001$) and univariate levels (NT: $F(1, 150) = 11.66$, $p < .01$; PF: $F(1, 150) = 13.99$, $p < .001$; A_{Exp}: $F(1, 150) = 9.82$, $p < .01$; PI: $F(1, 150) = 4.63$, $p < .05$) qualifies these main effects. Simple contrasts provide support for hypothesis 3c. The narrative reading participants who read the curiosity-order drama reported higher levels on the dependent measures than participants in any of the other three conditions (NT: mean differences ≥ .62, ps ≤ .001; PF: mean differences ≥ 1.56, ps ≤ .001; A_{Exp}: mean

differences ≥ .45, *p*s ≤ .05; PI: mean differences ≥ 4.03, *p*s ≤ .001; see table 5. In summary, when narrative reading is not inhibited, drama affects consumer responses, as demonstrated by the comment of a curiosity-order participant: “It was like a story. The reviewer didn't just post the review for accommodations in Agra, they described their journey from Delhi to Agra, giving some insights to their trip.”

Next, we bootstrapped the indirect effects of the instruction × drama interaction on positive feedback and purchase intention, using Hayes’s (2013a) models 4 and 6, respectively. The bootstrap estimates presented here are based on 500 bootstrap samples. Narrative transportation predicts the other dependent measures (PF: $\beta = .39$, $t = 2.43$, $p < .05$; A_{Exp}: $\beta = .41$, $t = 3.33$, $p < .01$; PI: $\beta = .67$, $t = 2.12$, $p < .05$) beyond the instruction × drama interaction. Narrative transportation and attitude toward the reviewed experience in serial mediate the relationship between the interaction and purchase intention (point estimate = .07, 95% CI = .02, .17). Positive feedback and attitude toward the reviewed experience in serial also mediate the relationship between the interaction and purchase intention (point estimate = .13, 95% CI = .05, .06). Finally, narrative transportation, positive feedback, and attitude toward the reviewed experience mediate the relationship between the instruction × drama interaction and purchase intention (point estimate = .03, 95% CI = .02, .17), in support of hypotheses 1a–1c. No other series are significant.

Alternative Explanation Test. We theorize that, more than the surprise order, the curiosity order stimulates drama, which in turn evokes narrative transportation more. Alternatively, we could argue that the curiosity order stimulates piecemeal processing more, because consumers attempt to piece together how the dramatic opening event came to pass, and this process

determines marketplace-related consumer responses instead. We disconfirmed this alternative. Participants' piecemeal processing did not differ between conditions ($ts_{(152)} \leq .43$, $ps \geq .672$).

CONCLUSION

The findings of our three studies provide several insights into the narrativity of experience reviews and the consequential consumer responses. They explain that textual, narrative elements affect narrative transportation, which explains positive feedback and helps predict attitude toward the reviewed experience and purchase intention (hypotheses 1a–1c). Overall, the significant narrative elements fall into three categories: (1) identifiable characters, (2) imaginable plot, and (3) genre. In the first category, we identify landscapes of affective and cognitive consciousness, whose presence, consistent with our predictions, influences consumer responses (hypotheses 2a and 2b). They clarify what the character is feeling and thinking.

In the second category, we identify temporal and spatial embedding and drama, whose presence, consistent with our predictions, influences consumer responses (hypotheses 3a–3c). They clarify the temporal and causal sequence and spatiality of events. We note that the effect size of temporal embedding is particularly large in study 1. The IRR shows that if a review's temporal embedding were to increase by .20 unigrams, positive feedback would increase by a factor of 1.22. This effect size highlights the importance of temporality and causality (Escalas, 1998, Woermann and Rokka, 2015). We propose that narrativity always indicates the existence of temporality and causality, and the creation of such features would be an act of storytelling. This simple definitional criterion would not impede reviewers from producing other narrative elements however. Although temporal embedding and other elements may coincide within the

same story, temporal embedding would precede in time and be conceptually distinct. We also note that study 3 provides a specific reason for the effect of drama—a story following a curiosity order of events.

In the third category, we identify comedy and tragedy, which, consistent with our prediction, influence consumer responses more than a progressive, regressive, or stable genre (hypothesis 4). Comedies and tragedies involve emotional story shapes that change over the course of a story line. These story shapes are more engaging than shapes that do not alternate in sign (Vonnegut, 2005). As such, we question the adequacy of the previously accepted negativity bias (Yin et al., 2014, Wu, 2013, Ludwig et al., 2013) by calling for its substitution with a new organizing proposition rather than merely claiming that the influence of review emotionality is completely incapable of being structured. We assert that review emotionality, which prior research considers a text-level phenomenon and a property of a review as a whole (Yin et al., 2014), is a sentence-level phenomenon and the property of the sentences that make up a review's emotional story shape. To test this assertion, we analyzed *post hoc* whether it matters for the study 1 comedies and tragedies where the respective trough or peak is. Complementing Cowley's (2014) work on peak and trough intensity on storytelling consumers, we note that the earlier the trough falls in the comedy, the more positive feedback the experience review garners from story-receiving consumers ($\beta = -.07$, SE = .09, $p < .05$, IRR = .93; see appendix G). Although this result does not generalize to the peak in tragedies ($\beta = .02$, SE = .10, $p = .521$), our newly developed tool to construct sentence-level emotional valence appears able to further explain the seemingly text-level negativity bias.

Regarding the unexpected control variables effects in study 1, the results show that pictures do not affect positive feedback. More precisely, the number of visitor photos per

experience does not influence consumer responses to the experience review. Pictures are static—as narrativity is important for experiential reviews' effect, pictures do not usually capture such. Indeed, the narrative element with which pictures correlates most strongly is spatial embedding ($\rho = .10$). Some pictures could be perceived as more dynamic, or multiple pictures could capture narrative elements other than spatial embedding (Senior et al., 2002).

Contributions

Our contributions to the narratology and WOM language literature streams are threefold. First, we provide a novel instrument for determining the development of narrativity within and across sentences and for examining whether there are intertextual differences in this. Adopting computational linguistics, we conduct an automated text analysis of *n*-grams of multiple word lengths. Moving beyond Chen and Berger's (2013) and Humphreys and Thompson's (2014) unigrammatic procedures, we organize the relationships among words in an online corpus of big data and assess identifiable characters and imaginable plot. To refine our understanding of genre, we further adapt growth-rate modeling to these *n*-grams (Jokisaari and Nurmi, 2009).

Second, we investigate the differences in positive feedback for experience reviews, which are narrative to different degrees. So far, research had not tied the evidence from narratology to the literature on positive feedback. This study establishes novel links between narrative elements and positive feedback. Specifically, we examine experience reviews and demonstrate that narrative elements related to (1) identifiable characters, (2) an imaginable plot, and (3) genres that involve changing emotional story shapes give consumers a transformational experience, which eventually affects positive feedback for these reviews.

Third, in addition to assessing experience reviews on the positive feedback potential they represent, we show what kind of reviews are more conducive to consumer behavior in the marketplace. Our research suggests that the more narrative a review is, the more transporting it is, in turn facilitating assessment for consumers. Specifically, more storied forms of reviews are more likely to change consumer attitudes and intentions than less storied ones.

Limitations and Directions for Further Research

As with any research, our studies suffer certain limitations. First, although our aim was to advance understanding of stories' impact on consumers, we specifically analyzed reviews of leisure travel–related experiential purchases. Several scholars (Scott and Orlikowski, 2012, van Laer et al., forthcoming) conceptualize travel as experience; thus, we build on substantial contributions to the field. However, the reviews we research address only the world's most purchased leisure travel–related experiences. This scale is sufficient to provide meaningful insight into the extent to which stories may have an impact on consumers, but the scope is not. Although people's motivations for and narrative interpretations of experience consumption are often surprisingly similar (Celsi et al., 1993), we call for additional investigation of the relationship between narrativity and consumer responses in other experience settings as well as for material purchases.

Second, we explain consumer responses to reviews using two narrative elements related to identifiable characters, three related to imaginable plot, and two related to genre; however, the case can always be made that other factors account for variance. Our text analysis operates at the sentence level. A sentence is the smallest set of words that is a complete unit of expression (e.g.,

of an event, a feeling, a thought, or an action) and that, in writing, must begin with a capital letter and conclude with a full stop, question mark, or exclamation mark. However, this important syntactic unit does not allow the inclusion of *n*-grams that exist within a sentence or genres that exist without an emotional story shape.

Third, our experimental studies show the need for research that goes beyond the limitations of computational linguistics. As another example, consider the assessment of genre. Gergen and Gergen (1988) were not the first to make the case that genre can account for variance in consumer responses. The earliest articulation of genre dates back to the Greek philosopher Plato (380BC/2008). Aristotle (335BC/1998) revised and extended the three Platonic genres (drama, dithyramb, and epic) to four: tragedy, epic, comedy, and parody. The number and types of genres have been modified many times since (Stern, 1995). For example, Genette (1979/1992) develops a structural poetic perspective on Plato's work and argues that the way he classified story genre relied on different modes of imitation of the world. We note that these taxonomies may constitute valid classifications of genres as well, but not ones that allow for the elaboration of specific hypotheses on story genre that are testable with computational linguistics. This methodology can integrate the narratology and WOM language literature streams in other ways however. We briefly discuss some avenues next.

(Re-)Conceptualizing "Story." We conceptualize a story by its narrative elements. This conceptualization can further expand research on narrativity. Pan and Zhang (2011) show that lengthy book reviews receive more positive feedback on Amazon.com. However, Hennig-Thurau, Wiertz, and Feldhaus (2015) report that consumers on Twitter consider brevity a virtue of movie reviews, and our data indicate that reviews on TripAdvisor can be very short. The increasingly popular flash fiction reinforces the notion of very short stories, such as in the case of

the alleged six-word novel *For Sale: Baby Shoes, Never Worn*. The research questions are numerous: "How short can a story be and still truly be a story?" (Thomas and Shapard, 2010, p. 12). Can consumers experience narrative transportation when reading these very short stories? If so, what is the effect of stories' length on consumer responses?

Effects on Conversion. Empirical work on the conversion effects of narratives is scant. Van Laer et al.'s (2014) review shows that only two storytelling studies measure actual behavior. However, recent developments in digital libraries indicate that there is ample opportunity to investigate conversion as an additional consequence of a narrative structure. We note that millions of e-books and fanfiction works are freely available on the Internet through Google Books, the Internet Archive, Project Gutenberg, Archive of Our Own, and Wattpad and propose a preliminary investigation of these texts in terms of the information they offer. Our text analytical tool may provide some explanation for the different conversion rates of these texts. Together with additional explanatory variables, such as Internet access, retail price, sales promotion, and WOM, these texts' narrativity may influence consumers' evaluations and decisions to download them. Research into these and other potential story-consistent conversion effects is clearly required.

Effects on Self-Branding. Arguably, economics would state that review content should meet an efficiency criterion: Consumers expect that reviewers give all information as efficiently as possible. From our text analysis, however, it appears that experience reviews are personal stories that encourage empathy. Under an economic explanation, such narrativity is unnecessary, does not meet the efficiency criterion, and should negatively affect consumer responses. Yet consumers who read such narratives do not seem convinced of the criterion's necessity; rather, they express a strong need for "the transfer of symbolic meaning of goods through storytelling"

(Arsel and Dobscha, 2011, p. 66). Such transfers could facilitate reviewer self-branding (Gandini, 2015), an asset that economics undermines or at least undervalues (McQuarrie et al., 2015). This argument marks a starting point from which to initiate research into possible personal brand creation or strengthening, because reviewers write narratives that violate economic principles.

Conclusion

Paraphrasing T. S. Eliot (1942), at the end of our exploration, we arrive where we started. Our approach led us to the observation that some experience reviews affect consumer responses within and beyond their hosting site, and we tried to understand why. To answer this research question, we used transportation theory and showed that the fuel of these influential reviews is the power of storytelling.

APPENDIX A

WORDS IN DICTIONARY ENTRY RELATED TO SURPRISE

amaz*
astonish*
astound*
awe
befuddl*
bewilder*
confound*
confus*
consternation
daze*
discomfiture
disconcert*
dismay*
dumbfound*
eye-open*
flabbergast*
jolt*
marvel*
perplex*
revelat*
shock*
startl*
stun
stunned
stunner
stupefy
surpris*
twist
wonder
wondered
wondering
wonders

The asterisk substitutes any other possible character(s) in the string.

APPENDIX B

OPERATIONALIZATION NARRATIVE ELEMENTS RELATED TO GENRE

Genres of changing emotional story shapes (i.e., the slope alternates in sign) are presumed to receive more helpfulness votes than those that have a progressive, regressive, or stable shape. Thus to measure a review's genre, we matched the words in each sentence against the LIWC dictionary entries relating to positive (406 words; e.g., love, nice, sweet) and negative emotions (499 words; e.g., hurt, nasty, ugly). We then converted the word counts into ratios of the number of words in the sentence, which represented the sentence-level "positive emotion" and "negative emotion" unigrams. We then calculated the sentence-level "emotion" *D*-gram (absolute difference between two *n*-grams). To investigate the emotional story shapes related to the different genres at the review-level, we estimated linear and nonlinear slopes. First, we numbered each sentence (*s*) from (1) opening sentence to (*k*) closing sentence. Second, we constructed a growth rate model with three variables—that is, an intercept, a linear degree variable and a curvilinear degree variable—estimating story shapes over the course of the review by regressing *s* and s^2 on the sentence-level emotion *D*-gram, using the method of least squares (Jokisaari and Nurmi 2009).

A significant, positive coefficient for *s* ($p < .05$) with or without a significant, positive coefficient for s^2 ($p > .05$) described the presence of a linear degree of increase of the emotion *D*-gram shape over the course of a review. We classified these reviews as progressive genres (n = 793). A significant, negative coefficient for *s* ($p < .05$) with or without a significant, negative coefficient for s^2 ($p > .05$) described the presence of a linear degree of decrease of the emotion *D*-gram shape over the course of a review. We classified these reviews as regressive genres (n = 4,601). No significant coefficient for *s* described a rate of change near zero for the emotion *D*-gram shape over the course of a review. We classified these reviews as stable genres (n = 163,576).

A significant, positive coefficient for *s* with a significant, negative coefficient for s^2 described the presence of a negative curvilinear degree of the emotion *D*-gram shape (i.e., u-shape) over the course of a review. We classified these reviews as comedies (n = 17,279). A significant, negative coefficient for *s* with a significant, positive coefficient for s^2 described the presence of a positive curvilinear degree of the emotion *D*-gram shape (i.e., inverted u-shape) over the course of a review. We classified these reviews as tragedies (n = 4,212).

APPENDIX C

VUONG TEST AND ZERO-INFLATED POISSON REGRESSION

To test whether negative binomial or zero-inflated Poisson regression provides an improvement over a more parsimonious, standard Poisson regression, we performed the test that Vuong (1989) proposes, using the Stata 14 software program. The test is bidirectional and tests the null hypotheses that neither of the two nonnested models (e.g., the zero-inflated Poisson and the standard Poisson regression) outperforms the other. The empirical values for Vuong's *Z* were above the critical values and positive, which disfavors the standard Poisson regression.

The dependent variable in the zero-inflated Poisson regression is specified as
$H_i \sim 0$ with probability q_i
$H_i \sim$ Poisson with probability $1 - q_i$,
where H_i represents the expected positive feedback received by review i.
We obtain q_i through

$$q_i = \frac{e^{\gamma' w_i}}{1 + e^{\gamma' w_i}}, \quad \textbf{(1)}$$

where γ' represents the parameter estimates β for the control variables (models 1–4), identifiable characters (models 2–4), imaginable plot (models 3 and 4), and genre (model 4) elements w_i that predict whether or not positive feedback is received. Conditional on receiving at least one count of positive feedback, positive feedback received follows a standard Poisson process,

$$H_i^* = e^{\beta' x_i}, \quad \textbf{(2)}$$

where x_i denote the control variables (models 1–4), identifiable characters (models 2–4), imaginable plot (models 3 and 4), and genre (model 4) elements with parameters β' for the conditional positive feedback counts H_i*.
We then generate H_i through

$$H_i = z_i H_i^*, \quad \textbf{(3)}$$

where z_i represents the (0/1) outcome of the binary model and H_i* is the Poisson-distributed positive feedback received given that $z_i = 1$.

APPENDIX D

SEPARATE ESTIMATION BINARY AND COUNT MODELS AS WELL AS MEANS AND STANDARD DEVIATIONS OF THE NARRATIVE ELEMENTS AS A FUNCTION OF CAST VOTE

	Binary Model		Count Model	
	β	(SE)	β	(SE)
Identifiable characters				
Landscape of affective consciousness	.11	(.01)***	.05	(.01)***
Landscape of cognitive consciousness	.03	(.01)***	.02	(.01)**
Imaginable plot				
Temporal embedding	.21	(.01)***	.08	(.01)***
Spatial embedding	.04	(.01)***	.02	(.01)***
Drama	-.01	(.01)	.00	(.01)
Genre				
Progressive	.18	(.09)*	.00	(.06)
Regressive	.06	(.04)	-.06	(.03)*
Comedy	.06	(.03)*	.04	(.02)*
Tragedy	.12	(.05)**	.08	(.03)**
Control variables[a]				
Experience rank order	-.06	(.09)	-.03	(.05)
Pictures	.11	(.10)	-.07	(.05)
Review age	.37	(.03)***	.26	(.01)***
Review eloquence	.02	(.01)	.02	(.01)***
Review extremity	.24	(.03)***	.10	(.01)***
Review readability	-.03	(.01)**	-.01	(.01)*
Review volume	-.70	(.07)***	-.02	(.04)
Reviewer expertise	.19	(.02)***	.04	(.00)***
McFadden's pseudo-R^2		.116		.149

[a] Though not reported in the table for the sake of brevity, we also included dummy variables in the model estimation to control for seemingly correlated reviews of the same experience and the 18 categories TripAdvisor uses to classify the multitude of experiences reviewed on its website.

Binary model: N = 190,461; count model: N = 66,187; * $p < .05$; ** $p < .01$; *** $p < .001$.

APPENDIX E

SENSITIVITY ANALYSIS

	Reduced Model		
	β	(SE)	IRR
Identifiable characters			
Landscape of affective consciousness	.07	(.01)***	1.07
Landscape of cognitive consciousness	.01	(.01)*	1.01
Imaginable plot			
Temporal embedding	.20	(.01)***	1.22
Spatial embedding	.05	(.01)**	1.05
Drama	-.02	(.02)	.98
Genre			
Progressive	.04	(.07)	1.04
Regressive	-.03	(.04)	.97
Comedy	.07	(.03)*	1.07
Tragedy	.08	(.04)*	1.09
Control variables[a]			
Experience rank order	-.06	(.08)	.94
Pictures	-.01	(.09)	.99
Review age	.35	(.02)***	1.42
Review eloquence	.04	(.01)**	1.04
Review extremity	.16	(.02)***	1.17
Review readability	-.03	(.01)**	.97
Review volume	-.48	(.09)***	.62
Reviewer expertise	.06	(.01)***	1.06
McFadden's pseudo-R^2		.150	

[a] Though not reported in the table for the sake of brevity, we also included dummy variables in the model estimation to control for seemingly correlated reviews of the same experience and the 18 categories TripAdvisor uses to classify the multitude of experiences reviewed on its website.

Reduced model: $N = 153{,}438$; * $p < .05$; ** $p < .01$; *** $p < .001$.

APPENDIX F

CURIOSITY AND SURPRISE EVENT ORDER

Curiosity Event Order

I did not get any sleep in Agra, home of the Taj Mahal. My holidays started out fine. After I visited the capital of India, Delhi, I moved on to see the Taj Mahal in Agra. Agra is only a short trip from Delhi. The Taj is a mausoleum built by Shah Jahan for his empress and is widely regarded as the most beautiful man-made structure in the world. It is said to be remarkable at all time of the day. It sure was when I visited as the sun rose above the early morning mists. Later that day, I went to look for accommodation in Agra. It turns out that Agra accommodations are not luxurious and I spent my nights awake on a straw mat.

Surprise Event Order

After I visited the capital of India, Delhi, I moved on to see the Taj Mahal in Agra. Agra is only a short trip from Delhi. The Taj is a mausoleum built by Shah Jahan for his empress and is widely regarded as the most beautiful man-made structure in the world. It is said to be remarkable at all time of the day. It sure was when I visited as the sun rose above the early morning mists. Later that day, I went to look for accommodation in Agra. Up until that moment, my holidays had been fine, but I did not get any sleep in Agra, home of the Taj Mahal. It turns out that Agra accommodations are not luxurious and I spent my nights awake on a straw mat.

APPENDIX G

COMEDY AND TRAGEDY: PEAK AND TROUGH LOCATION TEST

	β	Model 5 (SE)	IRR
Identifiable characters			
Landscape of affective consciousness	.06	(.01)***	1.06
Landscape of cognitive consciousness	.02	(.01)**	1.02
Imaginable plot			
Temporal embedding	.20	(.01)***	1.23
Spatial embedding	.05	(.02)***	1.05
Drama	-.02	(.02)	.98
Genre			
Progressive	-.12	(.07)	.89
Regressive	.01	(.05)	1.01
Comedy	.11	(.03)	1.07
Tragedy	.19	(.04)**	1.12
Control variables[a]			
Experience rank order	-.07	(.08)	.93
Pictures	.01	(.09)	1.01
Review age	.34	(.02)***	1.41
Review eloquence	.03	(.01)	1.03
Review extremity	.17	(.02)***	1.19
Review readability	-.03	(.01)***	.97
Review volume	-.50	(.10)***	.61
Reviewer expertise	.06	(.01)***	1.06
Peak and trough location			
Sentence number peak	-.07	(.09)**	.93
Sentence number trough	.02	(.10)	1.02
McFadden's pseudo-R^2		.154	

[a] Though not reported in the table for the sake of brevity, we also included dummy variables in the model estimation to control for seemingly correlated reviews of the same experience and the 18 categories TripAdvisor uses to classify the multitude of experiences reviewed on its website.

N = 85,413; reviews with less than three sentences or more than one peak or trough were excluded from the sample; * $p < .05$; ** $p < .01$; *** $p < .001$.

REFERENCES

ABELSON, R. P. 1981. Psychological status of the script concept. *American Psychologist,* 36**,** 715-729.

ADAVAL, R., ISBELL, L. M. & WYER, R. S. 2007. The impact of pictures on narrative- and list-based impression formation: A process interference model. *Journal of Experimental Social Psychology,* 43**,** 352-364.

ADAVAL, R. & WYER, R. S. 1998. The role of narratives in consumer information processing. *Journal of Consumer Psychology,* 7**,** 207-245.

ALBA, J. W. & HASHER, L. 1983. Is memory schematic? *Psychological Bulletin,* 93**,** 203-231.

ANDERSON, E. T. & SIMESTER, D. I. 2014. Reviews without a purchase: Low ratings, loyal customers, and deception. *Journal of Marketing Research,* 51**,** 249-269.

ARGO, J. J., ZHU, R. & DAHL, D. W. 2008. Fact or fiction: An investigation of empathy differences in response to emotional melodramatic entertainment. *Journal of Consumer Research,* 34**,** 614-623.

ARISTOTLE 335BC/1998. *Poetics,* London, Nick Hern.

ARSEL, Z. & DOBSCHA, S. 2011. Hybrid pro-social exchange systems: The case of freecycle. *In:* AHLUWALIA, R., CHARTRAND, T. L. & RATNER, R. K. (eds.) *Advances in Consumer Research.* Duluth, MN: Association for Consumer Research.

BARGH, J. A. & CHARTRAND, T. L. 2000. The mind in the middle: A practical guide to priming and automaticity research. *In:* REIS, H. T. & JUDD, C. M. (eds.) *Handbook of research methods in social and personality psychology.* Cambridge: Cambridge University.

BARNHARDT, T. M., MANZANO, I., BRITO, M., MYRICK, M. & SMITH, S. M. 2016. The effects of product placement in fictitious literature on consumer purchase intention. *Psychology & Marketing,* 33**,** 883-898.

BARTHES, R. 1975. An introduction to the structural analysis of narrative. *New Literary History,* 6**,** 237-272.

BAUMEISTER, R. F. & NEWMAN, L. S. 1994. How stories make sense of personal experiences: Motives that shape autobiographical narratives. *Personality and Social Psychology Bulletin,* 20**,** 676-690.

BENNETT, A. & ROYLE, N. 2004. *Introduction to literature, criticism and theory,* Harlow, Pearson.

BERGER, J. 2014. Word of mouth and interpersonal communication: A review and directions for future research. *Journal of Consumer Psychology,* 24**,** 586-607.

BERGER, J. & IYENGAR, R. 2013. Communication channels and word of mouth: How the medium shapes the message. *Journal of Consumer Research,* 40**,** 567-579.

BERTHOLD, M. R., CEBRON, N., DILL, F., GABRIEL, T. R., KÖTTER, T., MEINL, T., OHL, P., SIEB, C., THIEL, K. & WISWEDEL, B. 2008. KNIME: The KoNstanz Information MinEr. *In:* PREISACH, C., BURKHARDT, H., SCHMIDT-THIEME, L. & DECKER, R. (eds.) *Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V., Albert-Ludwigs-Universität Freiburg, March 7–9, 2007.* Berlin, Heidelberg: Springer Berlin Heidelberg.

BREWER, W. F. & LICHTENSTEIN, E. H. 1981. Event schemas, story schemas, and story grammars. *In:* LONG, J., BADDELEY, A. & BARNARD, P. (eds.) *Attention and Performance IX : 9th international symposium, Cambridge, July 1980, proceedings.* Hillsdale, NJ: Lawrence Erlbaum.

BREWER, W. F. & LICHTENSTEIN, E. H. 1982. Stories are to entertain: A structural-affect theory of stories. *Journal of Pragmatics,* 6**,** 473-486.

BRUNER, J. S. 1986. *Actual minds, possible worlds,* Cambridge, MA, Harvard.

BURKE, K. 1962. *A grammar of motives and a rhetoric of motives* Cleveland, OH, World Publishing.

CAO, Q., DUAN, W. & GAN, Q. 2011. Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. *Decision Support Systems,* 50**,** 511-521.

CELSI, R. L., ROSE, R. L. & LEIGH, T. W. 1993. An exploration of high-risk leisure consumption through skydiving. *Journal of Consumer Research,* 20**,** 1-23.

CHEN, Z. & BERGER, J. 2013. When, why, and how controversy causes conversation. *Journal of Consumer Research,* 40**,** 580-593.

CHEN, Z. & BERGER, J. 2016. How content acquisition affects word of mouth. *Journal of Consumer Research,* 43**,** 86-102.

COWLEY, E. 2014. Consumers telling consumption stories: Word-of-mouth and retrospective evaluations. *Journal of Business Research,* 67**,** 1522-1529.

DAS, S. R. & CHEN, M. Y. 2007. Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management Science,* 53**,** 1375-1388.

DZHOGLEVA, H. & LAMBERTON, C. P. 2014. Should birds of a feather flock together? Understanding self-control decisions in dyads. *Journal of Consumer Research,* 41**,** 361-380.

ELIOT, E. & ANDERSON, M. 2015. Local Consumer Review Survey. Lewes, UK.

ELIOT, T. S. 1942. *Little Gidding,* London, Faber and Faber.

ESCALAS, J. E. 1998. Advertising narratives. What are they and how do they work? *In:* STERN, B. B. (ed.) *Representing consumers: Voices, views, and visions.* London: Routledge.

ESCALAS, J. E. 2004. Imagine yourself in the product: Mental simulation, narrative transportation, and persuasion. *Journal of Advertising,* 33**,** 37-48.

ESCALAS, J. E. 2007. Self-referencing and persuasion: Narrative transportation versus analytical elaboration. *Journal of Consumer Research,* 33**,** 421-429.

ESCALAS, J. E. & BETTMAN, J. R. 2000. Using narratives to discern: Self-identity related consumer goals and motivations. *In:* RATNESHWAR, R., MICK, D. G. & HUFFMAN, C. (eds.) *The why of consumption: Perspectives on consumer motives, goals, and desires.* New York, NY: Routledge.

FELDMAN, C. F., BRUNER, J. S., RENDERER, B. & SPITZER, S. 2014. Narrative comprehension. *In:* BRITTON, B. K. & PELLEGRINI, A. D. (eds.) *Narrative thought and narrative language.* New York, NY: Taylor & Francis.

FIGUEIREDO, B. & SCARABOTO, D. 2016. The systemic creation of value through circulation in collaborative consumer networks. *Journal of Consumer Research,* 43**,** 509-533.

FORMAN, C., GHOSE, A. & WIESENFELD, B. 2008. Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research,* 19**,** 291-313.

GANDINI, A. 2015. Digital work: Self-branding and social capital in the freelance knowledge economy. *Marketing Theory*.

GENETTE, G. 1979/1992. *The architext: An introduction,* Berkeley, CA, University of California.

GERGEN, K. J. & GERGEN, M. M. 1988. Narrative and the self as relationship. *In:* BERKOWITZ, L. (ed.) *Advances in Experimental Social Psychology.* New York, NY: Academic Press.

GERRIG, R. J. 1993. *Experiencing narrative worlds: On the psychological activities of reading,* New Haven, CT, Yale.

GHOSE, A. & IPEIROTIS, P. G. 2011. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Transactions on Knowledge & Data Engineering,* 23**,** 1498-1512.

GODES, D. & MAYZLIN, D. 2004. Using online conversations to study word-of-mouth communication. *Marketing Science,* 23**,** 545-560.

GREEN, M. C. 2004. Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes,* 38**,** 247-267.

GREEN, M. C. 2008. Transportation theory. *In:* DONSBACH, W. (ed.) *International Encyclopedia of Communication.* Oxford, UK: Wiley-Blackwell.

GREEN, M. C. & BROCK, T. C. 2000. The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology,* 79**,** 701-721.

GREEN, M. C. & BROCK, T. C. 2002. In the mind's eye: Transportation-imagery model of narrative persuasion. *In:* GREEN, M. C., STRANGE, J. J. & BROCK, T. C. (eds.) *Narrative impact: Social and cognitive foundations.* Mahwah, NJ: Lawrence Erlbaum.

GREENE, W. H. 2011. *Econometric Analysis,* Upper Saddle River, NJ, Pearson.

HAYES, A. F. 2013a. *Introduction to Mediation, Moderation, and Conditional Process Analysis,* New York, NY, Guilford.

HAYES, A. F. Multilevel mediation analysis. Association for Psychological Science Conference, 23 May 2013b Washington, DC.

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? *Journal of Interactive Marketing,* 18**,** 38.

HENNIG-THURAU, T., WIERTZ, C. & FELDHAUS, F. 2015. Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marketing Science,* 43**,** 375-394.

HOLBROOK, M. B. & GRAYSON, M. W. 1986. The semiology of cinematic consumption: Symbolic consumer behavior in Out of Africa. *Journal of Consumer Research,* 13**,** 374-381.

HUMPHREYS, A. & THOMPSON, C. J. 2014. Branding disaster: Reestablishing trust through the ideological containment of systemic risk anxieties. *Journal of Consumer Research,* 41**,** 877-910.

JOKISAARI, M. & NURMI, J.-E. 2009. Change in newcomers' supervisor support and socialization outcomes after organizational entry. *Academy of Management Journal,* 52**,** 527-544.

JURAFSKY, D., CHAHUNEAU, V., ROUTLEDGE, B. R. & SMITH, N. A. 2014. Narrative framing of consumer sentiment in online restaurant reviews. *First Monday,* 19.

JUSTER, F. T. 1966. Consumer buying intentions and purchase probability: An experiment in survey design. *Journal of the American Statistical Association,* 61**,** 658-696.

KRIPPENDORFF, K. 2013. *Content Analysis: An Introduction to Its Methodology,* Thousand Oaks, CA, Sage.

KRONROD, A. & DANZIGER, S. 2013. "Wii will rock you!" The use and effect of figurative language in consumer reviews of hedonic and utilitarian consumption. *Journal of Consumer Research,* 40**,** 726-739.

LAS VEGAS CONVENTION AND VISITORS AUTHORITY. 2017. *2016 Las Vegas Year-to-Date Executive Summary* [Online]. Las Vegas, NV. Available: http://www.lvcva.com/includes/content/images/media/docs/ES-YTD-2016.pdf [Accessed].

LEVY, S. J. 2006. The consumption of stories. *In:* RUSSELL W. BELK (ed.) *Handbook of Qualitative Research Methods in Marketing.* Cheltenham: Edward Elgar.

LOVE HOME SWAP. 2015. *The world's 50 most visited tourist attractions: Every travel destination has at least one must-see attraction* [Online]. Available: http://www.lovehomeswap.com/blog/latest-news/the-50-most-visited-tourist-attractions-in-the-world [Accessed 1 January].

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. *Journal of Marketing,* 77**,** 87-103.

MANDLER, J. M. 1984. *Stories, scripts, and scenes: Aspects of schema theory,* Hillsdale, NJ, Lawrence Erlbaum.

MARSH, E. J. & FAZIO, L. K. 2006. Learning errors from fiction: Difficulties in reducing reliance on fictional stories. *Memory & Cognition,* 34**,** 1140-1149.

MCENERY, T. & HARDIE, A. 2011. *Corpus linguistics: Method, theory and practice,* Cambridge, Cambridge University.

MCQUARRIE, E. F., MCINTYRE, S. H. & SHANMUGAM, R. 2015. Yelp and the soapbox imperative. *In:* MCQUARRIE, E. F. (ed.) *The new consumer online: A sociology of taste, audience, and publics.* Cheltenham: Edward Elgar.

MERRIAM-WEBSTER. 2015. *Surprise Synonyms | Merriam-Webster Thesaurus* [Online]. Available: http://www.merriam-webster.com/thesaurus/surprise [Accessed].

MOORE, S. G. 2012. Some things are better left unsaid: How Word of Mouth influences the storyteller. *Journal of Consumer Research,* 38**,** 1140-1154.

MOORE, S. G. 2015. Attitude predictability and helpfulness in online reviews: The role of explained actions and reactions. *Journal of Consumer Research,* 42**,** 30-44.

MUDAMBI, S. M. & SCHUFF, D. 2010. What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*, 34**,** 185-200.

NIELSEN 2015. *Global trust in advertising: Winning strategies for an evolving media landscape*, New York, NY.

NIELSEN, J. H. & ESCALAS, J. E. 2010. Easier is not always better: The moderating role of processing type on preference fluency. *Journal of Consumer Psychology*, 20**,** 295-305.

OSGOOD, C. E., SUCI, G. J. & TANNENBAUM, P. H. 1957. *The measurement of meaning*, Chicago, IL, University of Illinois.

OXFORD DICTIONARIES. 2015. *surprise: definition of surprise in Oxford Dictionaries (Thesaurus of English)* [Online]. Available: http://www.oxforddictionaries.com/definition/english-thesaurus/surprise [Accessed].

PAN, Y. & ZHANG, J. Q. 2011. Born unequal: A study of the helpfulness of user-generated product reviews. *Journal of Retailing*, 87**,** 598-612.

PECHMANN, C. & WANG, L. 2010. Effects of Indirectly and Directly Competing Reference Group Messages and Persuasion Knowledge: Implications for Educational Placements. *Journal of Marketing Research*, 47**,** 134-145.

PENNEBAKER, J. W., BOYD, R. L., JORDAN, K. & BLACKBURN, K. 2015. The Development and Psychometric Properties of LIWC2015. Austin, TX: University of Texas at Austin.

PENNEBAKER, J. W., CHUNG, C. K., IRELAND, M., GONZALES, A. & BOOTH, R. J. 2007. The Development and Psychometric Properties of LIWC2007. Austin, TX: LIWC.net.

PHILLIPS, B. J. & MCQUARRIE, E. F. 2010. Narrative and persuasion in fashion advertising. *Journal of Consumer Research*, 37**,** 368-392.

PLATO 380BC/2008. *Republic*, Oxford, Oxford University.

PORTER ABBOTT, H. 2002. *The Cambridge Introduction to Narrative*, Cambridge, Cambridge University.

RUSSELL, C. A. 2002. Investigating the effectiveness of product placements in television shows: The role of modality and plot connection congruence on brand memory and attitude. *Journal of Consumer Research*, 29**,** 306-318.

SCHELLEKENS, G. A. C., VERLEGH, P. W. J. & SMIDTS, A. 2010. Language abstraction in word of mouth. *Journal of Consumer Research*, 37**,** 207-223.

SCHLOSSER, A. E. 2003. Experiencing products in the virtual world: The role of goal and imagery in influencing attitudes versus purchase intentions. *Journal of Consumer Research,* 30**,** 184-198.

SCOTT, S. V. & ORLIKOWSKI, W. J. 2012. Reconfiguring relations of accountability: Materialization of social media in the travel sector. *Accounting, Organizations and Society,* 37**,** 26-40.

SENIOR, C., WARD, J. & DAVID, A. S. 2002. Representational momentum and the brain: An investigation into the functional necessity of V5/MT. *Visual Cognition,* 9**,** 81-92.

SLATER, M. D. & ROUNER, D. 2002. Entertainment-education and elaboration likelihood: Understanding the processing of narrative persuasion. *Communication Theory,* 12**,** 117-244.

SOMERS, M. R. 1994. The narrative constitution of identity: A relational and network approach. *Theory and Society,* 23**,** 605-649.

STERN, B. B. 1995. Consumer myths: Frye's taxonomy and the structural analysis of consumption text. *Journal of Consumer Research,* 22**,** 165-185.

STURGESS, P. J. M. 1992. *Narrativity: Theory and practice,* Oxford, Clarendon.

TAUSCZIK, Y. R. & PENNEBAKER, J. W. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology,* 29**,** 24-54.

THOMAS, J. & SHAPARD, R. (eds.) 2010. *Flash Fiction Forward: 80 Very Short Stories,* New York, NY: Norton.

THOMPSON, C. J. 1997. Interpreting consumers: A hermeneutical framework for deriving marketing insights from the texts of consumers' consumption stories. *Journal of Marketing Research,* 34**,** 438-455.

TRIPADVISOR. 2017. *Read Reviews, Compare Prices & Book* [Online]. Available: http://www.tripadvisor.com/ [Accessed].

TRIPADVISOR SUPPORT. 2017. *What is TripAdvisor's Popularity Index?* [Online]. Available: https://www.tripadvisorsupport.com/hc/en-us/articles/200613987-What-is-TripAdvisor-s-Popularity-Index- [Accessed].

VAN BOVEN, L. & GILOVICH, T. 2003. To do or to have? That is the question. *Journal of Personality and Social Psychology,* 85**,** 1193-1202.

VAN DEN HENDE, E. A., DAHL, D. W., SCHOORMANS, J. P. L. & SNELDERS, D. 2012. Narrative transportation in concept tests for really new products: The moderating effect of reader–protagonist similarity. *Journal of Product Innovation Management,* 29**,** 157-170.

VAN LAER, T. & DE RUYTER, K. 2010. In stories we trust: How narrative apologies provide cover for competitive vulnerability after integrity-violating blog posts. *International Journal of Research in Marketing,* 27**,** 164-174.

VAN LAER, T., DE RUYTER, K., VISCONTI, L. M. & WETZELS, M. 2014. The Extended Transportation-Imagery Model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research,* 40**,** 797-817.

VAN LAER, T., VISCONTI, L. M. & FEIEREISEN, S. forthcoming. Need for narrative. *Journal of Marketing Management*.

VÁSQUEZ, C. 2014. *The discourse of online consumer reviews,* London, Bloomsbury.

VONNEGUT, K. 2005. *A man without a country: A memoir of life in George W Bush's America,* London, Bloomsbury.

WANG, J. & CALDER, B. J. 2006. Media transportation and advertising. *Journal of Consumer Research,* 33**,** 151-162.

WEBER, K. 2005. A toolkit for analyzing corporate cultural toolkits. *Poetics,* 33**,** 227-252.

WOERMANN, N. & ROKKA, J. 2015. Timeflow: How consumption practices shape consumers' temporal experiences. *Journal of Consumer Research,* 41**,** 1486-1508.

WOODSIDE, A. G., SOOD, S. & MILLER, K. E. 2008. When consumers and brands talk: Storytelling theory and research in psychology and marketing. *Psychology & Marketing,* 25**,** 97-145.

WU, P. F. 2013. In search of negativity bias: An empirical study of perceived helpfulness of online reviews. *Psychology & Marketing,* 30**,** 971-984.

YELP. 2017. *User Reviews and Recommendations of Top Restaurants, Shopping, Nightlife, Entertainment, Services and More* [Online]. Available: http://www.yelp.com/ [Accessed].

YIN, D., BOND, S. D. & ZHANG, H. 2014. Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews. *MIS Quarterly,* 38**,** 539-560.

TABLE 1

NARRATIVE ELEMENTS: DEFINITIONS, REPRESENTATIVE ARTICLES, AND STUDY 1 OPERATIONALIZATIONS

Element	Definition	Representative Articles	Operationalization
Identifiable characters			
Landscape of affective consciousness	The extent to which the review recounts an initial event about which a character expresses feelings that, in turn, lead to a course of action	Bruner (1986); Feldman et al. (2014)	Number of motion–affective process–motion trigrams divided by number of sentences in a review
Landscape of cognitive consciousness	The extent to which the review recounts an event about which a character expresses thoughts that, in turn, lead to a course of action	Bruner (1986); Feldman et al. (2014)	Number of motion–insight–motion trigrams divided by number of sentences in a review
Imaginable plot			
Temporal embedding	The extent to which the review is organized in a temporal sequence and provides causal links between the events that occur	Adaval et al. (2007); Adaval and Wyer (1998); Barthes (1975); Brewer and Lichtenstein (1981); Mandler (1984); Thompson (1997)	Presence of time unigrams (1), causation unigrams (2), time and causation unigrams (3)
Spatial embedding	The extent to which the review focuses on particular events	Baumeister and Newman (1994); Brewer and Lichtenstein (1981); Gerrig (1993)	Proportion of space unigrams to other words in a review
Drama	The extent to which the dramatistic	Burke (1962)	Proportion of surprise unigrams to other words in a review

	pentad is breached		
Genre		Gergen and Gergen (1988); Vonnegut (2005)	
Progressive	Emotion ameliorates over the course of the review		Continuous increase of emotional story shape
Regressive	Emotion deteriorates over the course of the review		Continuous decrease of emotional story shape
Stable	Emotion is stable over the course of the review		Rate of change near zero for emotional story shape
Comedy	Emotion first deteriorates and then ameliorates over the course of the review.		Negative curvilinear degree of emotional story shape (i.e., U shape)
Tragedy	Emotion first ameliorates and then deteriorates over the course of the review.		Positive curvilinear degree of emotional story shape (i.e., inverted U shape)

TABLE 2

STUDY 1: MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF POSITIVE FEEDBACK, NARRATIVE ELEMENTS, AND CONTROL VARIABLES

		M	(SD)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.	Positive feedback	.77	(2.01)																	
Identifiable characters																				
2.	Landscape of affective consciousness	.24	(.65)	.13																
3.	Landscape of cognitive consciousness	.08	(.34)	.11	.60															
Imaginable plot																				
4.	Temporal embedding	3.17	(.90)	.12	.20	.16														
5.	Spatial embedding	.06	(.04)	.03	.07	.04	.04													
6.	Drama	.07	(.03)	-.03	-.03	-.03	-.05	-.04												
Genre[a]																				
7.	Progressive	.00	(.06)	.02	.02	.02	.03	.00	-.01											
8.	Regressive	.02	(.15)	.00	.03	.02	.04	.01	.02	-.01										
9.	Comedy	.09	(.29)	.02	.06	.06	.08	.02	.02	-.02	-.05									

10.	Tragedy	.02	(.15)	.03	.03	.02	.03	-.01	-.02	-.01	-.02	-.05							
Control variables																			
11.	Experience rank order	.84	(.18)	-.16	-.03	-.03	-.07	-.04	.07	-.03	.01	.00	-.04						
12.	Pictures	545	(81)	-.11	-.02	-.02	-.09	.10	.05	-.02	-.02	-.04	-.03	.35					
13.	Review age	740	(569)	.38	.05	.04	.04	.01	-.02	.01	.00	.01	.01	-.13	-.01				
14.	Review eloquence	.14	(.06)	-.01	-.03	-.01	-.09	-.13	.05	.00	-.01	-.02	.00	.05	.01	-.02			
15.	Review extremity	.66	(.65)	.21	.03	.04	.07	.03	-.06	.07	-.02	-.04	.07	-.36	-.18	.13	-.05		
16.	Review readability	.00	(1.00)	-.03	-.01	-.02	-.06	-.04	.00	.00	-.01	-.01	.00	.03	.01	-.02	.01	-.05	
17.	Review volume	4,152	(7,274)	-.12	-.11	-.08	-.12	.01	.07	-.02	-.03	-.05	-.03	.33	.86	.00	.03	-.16	.01
	Reviewer expertise	27	(83)	.11	.02	.01	.03	.04	-.01	.00	.00	.01	.00	-.02	.01	.07	.00	.01	.00	.00

[a] Stable genre is the reference level.

Statistically significant correlations at $p < .05$ at the two-tailed level: $\rho \leq -.01$ or $\rho \geq .01$.

TABLE 3

STUDY 1: EFFECTS OF NARRATIVE ELEMENTS AND CONTROL VARIABLES ON POSITIVE FEEDBACK

	Model 1		Model 2		Model 3		Model 4		
	β	(SE)	β	(SE)	β	(SE)	β	(SE)	IRR
Identifiable characters									
Landscape of affective consciousness			.08	(.01)***	.07	(.01)***	.07	(.01)***	1.07
Landscape of cognitive consciousness			.02	(.01)***	.02	(.01)**	.02	(.01)**	1.02
Imaginable plot									
Temporal embedding					.20	(.01)***	.20	(.01)***	1.22
Spatial embedding					.04	(.01)***	.04	(.01)***	1.05
Drama					-.02	(.01)	-.02	(.01)	.98
Genre									
Progressive							.02	(.07)	1.02
Regressive							-.05	(.03)	.95
Comedy							.07	(.03)***	1.07
Tragedy							.10	(.04)**	1.10
Control variables[a]									
Experience rank order	-.06	(.08)	-.06	(.08)	-.06	(.08)	-.06	(.08)	.94
Pictures	.05	(.09)	.01	(.09)	.00	(.09)	.00	(.09)	1.00
Review age	.35	(.02) **	.35	(.02)***	.35	(.02)***	.35	(.02)***	1.42
Review eloquence	.01	(.01)	.01	(.01)	.03	(.01) **	.03	(.01) **	1.03
Review extremity	.17	(.02) **	.17	(.02)***	.16	(.02)***	.16	(.02)***	1.17
Review readability	-.04	(.01) **	-.04	(.01)***	-.03	(.01)***	-.03	(.01)***	.97
Review volume	-.57	(.09) **	-.53	(.09)***	-.51	(.09)***	-.51	(.09)***	.60
Reviewer expertise	.06	(.00) **	.06	(.00)***	.06	(.00)***	.06	(.00)***	1.06
Wald's $\chi^2_{\text{Change (df)}}$			2630.32$_{(2)}$ ***		3294.72$_{(3)}$ ***		97.07$_{(4)}$ ***		
McFadden's pseudo-R^2		.142		.147		.153		.154	

[a] Though not reported in the table for the sake of brevity, we also included dummy variables in the model estimation to control for seemingly correlated reviews of the same experience and the 18 categories TripAdvisor uses to classify the multitude of experiences reviewed on its website. Hypothesized effects that study 1 supported appear in bold.
All models: N = 190,461; Model 1: Wald's $\chi^2_{(27)}$ = 20759.15; * $p < .05$; ** $p < .01$; *** $p < .001$.

TABLE 4

STUDY 2: MEANS, STANDARD DEVIATIONS, INTERCORRELATIONS, AND BOOTSTRAP RESULTS OF POSITIVE FEEDBACK, NARRATIVE TRANSPORTATION, AND NARRATIVE ELEMENTS

	M	(SD)	1	2	Point Estimate	95% CI
1. Narrative transportation	3.84	(.94)				
2. Positive feedback	4.76	(1.71)	.56			
Identifiable characters						
Landscape of affective consciousness	.04	(.07)	.05	.06	.04	.01–.07
Landscape of cognitive consciousness	.02	(.05)	.06	.07	.06	.03–.09
Imaginable plot						
Temporal embedding	3.31	(.77)	.07	.07	.09	.05–.12
Spatial embedding	.05	(.04)	.17	.20	.20	.16–.23
Drama	.07	(.03)	.02	.05	.02	-.01–.06
Genre[a]						
Progressive	.03	(.18)	.06	.02		
Regressive	.05	(.21)	-.01	.01		
Comedy	.18	(.38)	.09	.11	.11	.08–.14
Tragedy	.17	(.37)	.04	.09	.04	.01–.06

[a] Stable genre is the reference level.

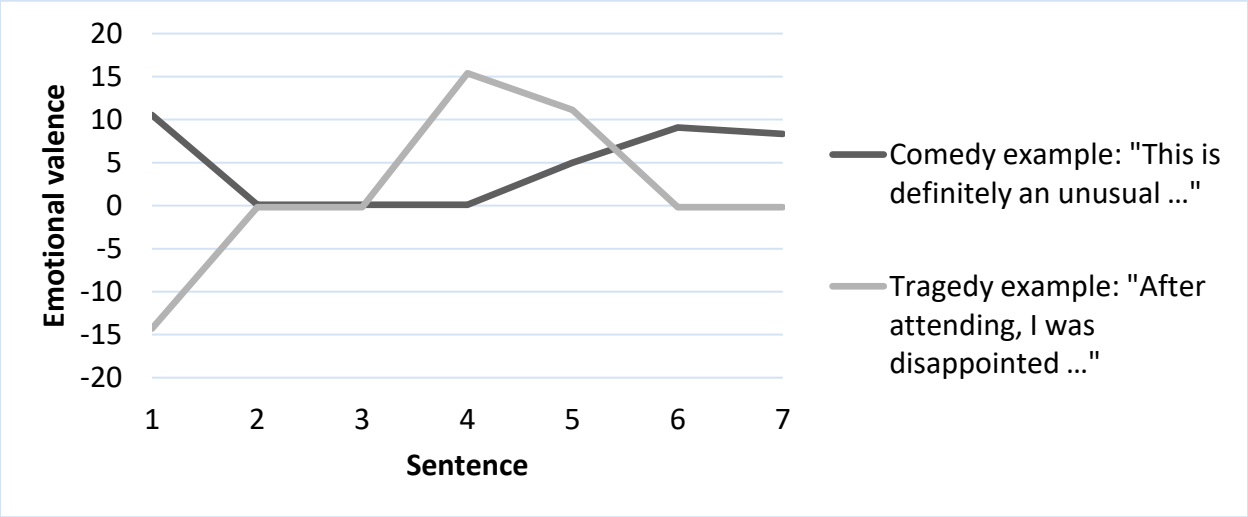
Hypothesized effects that study 2 support appear in bold. Statistically significant correlations at $p < .05$ at the two-tailed level: $\rho \leq -.04$ or $\rho \geq .04$. Statistically significant mediation at $p < .05$: 95% CI < .00 or 95% CI > .00.

TABLE 5

STUDY 3: MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS OF THE DEPENDENT VARIABLES AS A FUNCTION OF NARRATIVE AND AGE-10 READING INSTRUCTION AND CURIOSITY- AND SURPRISE-ORDER DRAMA

		Narrative				Age-10						
		Curiosity		Surprise		Curiosity		Surprise				
		N = 154		*n* = 43		*n* = 37		*n* = 38		*n* = 36		
		M	(SD)	*M*	(SD)	*M*	(SD)	*M*	(SD)	*M*	(SD)	1	2	3
1.	Narrative transportation	4.38	(.66)	4.94	(.38)	4.07	(.71)	4.32	(.43)	4.06	(.63)			
2.	Positive feedback	3.98	(1.35)	5.16	(.98)	3.56	(1.28)	3.61	(1.03)	3.38	(1.27)	.35		
3.	Attitude toward the reviewed experience	4.52	(.96)	5.16	(.52)	3.86	(1.10)	4.70	(.69)	4.26	(.96)	.34	.49	
	Purchase intention	5.16	(2.54)	6.98	(2.13)	4.28	(2.51)	5.09	(2.21)	3.95	(2.15)	.31	.42	.48

Statistically significant correlations at $p < .01$ at the two-tailed level: $\rho \geq .31$.

FIGURE 1

COMEDY AND TRAGEDY: EXAMPLE STORY SHAPES