This is the accepted and peer reviewed version of: Becker, J.-M., Völckner, F., & Sattler, H. (2024). How Important Is Word of Mouth? Development, Validation, and Application of a Scale. Journal of Interactive Marketing, 0(0). https://doi.org/10.1177/10949968231215362

How Important is Word-of-Mouth?

Development, Validation, and Application of a Scale

Author Accepted Manuscript

To appear in: *Journal of Interactive Marketing*

Accepted Date:2 November 2023

DOI: https://doi.org/10.1177/10949968231215362

Jan-Michael Becker*
BI Norwegian Business School, Norway,
jan-michael.becker@bi.no, https://orcid.org/0000-0003-3013-3739

Franziska Völckner
University of Cologne, Germany,
voelckner@wiso.uni-koeln.de, https://orcid.org/0000-0003-4835-6199

Henrik Sattler
University of Hamburg, Germany,
Henrik.Sattler@uni-hamburg.de,

* corresponding author

Acknowledgments: The authors thank Valentyna Melnyk and the participants of research seminars at KU Leuven and Massey University Albany for their helpful comments on previous versions of this article. The authors also thank Julia Pfeil and Birte Terlinden for their help in collecting the data for Studies 1 and 2.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the Deutsche Forschungsgemeinschaft (grant number VO 1555/1-1).

How Important is Word-of-Mouth? Development, Validation, and Application of a Scale

Abstract

Companies spend large amounts of money to induce word of mouth (WOM) and spread it among consumers. This research introduces the concept of "WOM relevance," which measures the importance of WOM for consumers' purchase-decision process in a specific category. It uses three studies to develop and validate a parsimonious scale to measure WOM relevance at the consumer level across various product categories and different types of WOM, and applies the scale in an additional set of five studies. Specifically, this research disentangles the consumer-level and category components of WOM relevance; shows that the consumer-level variation is (much) larger than the category-level variation; and provides insights into differences in WOM relevance across categories, consumers, and WOM types. It also empirically shows that electronic WOM relevance relates to consumers' search behavior and consideration-set formation in an online-shopping environment. Furthermore, it demonstrates that the proposed scale predicts choices as well as a more sophisticated choice model. Finally, this research shows that WOM relevance influences not only consumers' own purchase-decision process but also their intentions to retransmit others' WOM messages.

Keywords: (Electronic) word of mouth; Social media; Word-of-mouth relevance; Cross-category research; WOM retransmission; Consumer search; Scale development

Introduction

Word of mouth (WOM), defined as "informal [and experience-based (Houston et al. 2018)] communications directed at other consumers about the ownership, usage, or characteristics of particular goods and services or their sellers" (Westbrook 1987, p. 261), is ubiquitous in today's marketing environment. Companies spend large amounts of money to induce WOM and spread it among consumers to influence their decision making and stimulate sales because they believe it is one of the most important information sources for consumer decision making (Babić Rosario, de Valck, and Sotgiu 2020; Todorov 2021). Accordingly, Hewett et al. (2016) note that assessing the relevance of WOM to manage WOM-related marketing activities, create budget-allocation plans, and target specific consumer segments is a significant challenge managers face. This notion is also reflected in a small survey of management consultants that we conducted and report in the section "How Marketing Managers Think About WOM Relevance."

Academic studies show that WOM influences consumer decision making and subsequently product sales (e.g., Liu 2006; Zhu and Zhang 2010). In addition, meta-analyses indicate that the role of WOM may differ across category types (e.g., durable goods versus non-durable goods; You et al. 2015) and WOM types (e.g., consumer reviews versus social media posts; Babić Rosario et al. 2016). However, the underlying studies only consider WOM relevance at the aggregate market level (i.e., the sales impact of observable WOM characteristics, such as the volume, valence, and variance of WOM) and only of online channels. Importantly, these studies cannot identify consumer-level differences in WOM relevance, although there is some evidence that the relevance of WOM may vary across consumers. For example, Moliner-Velázquez, Fuentes-Blasco, and Gil-Saura's (2021) study reveals heterogeneity in consumer intentions to look at online reviews for hotels. Such consumer-level differences (i.e., consumer characteristics that increase or decrease WOM relevance) are particularly important when managers want to profile consumer segments for which certain WOM types are most relevant. These market-level studies also cannot identify the relevance of traditional offline WOM, which is more difficult to observe than online WOM. Furthermore, in terms of category characteristics, existing meta-analyses may suffer from selection bias because the underlying empirical studies typically focus on categories wherein WOM is highly prevalent. Thus, it remains unclear how relevant WOM is for consumers' purchase-decision process in categories wherein WOM is less prevalent.

Despite significant academic interest in understanding WOM (Verma and Yadav 2021), there is still no uniform instrument available to systematically measure the relevance of WOM at the consumer level across product categories and WOM types. Such an instrument would enable in-depth empirical comparison of the relevance of WOM within and across categories, consumers, and WOM types, complementing data mining–based web metrics that track WOM availability, content, and sentiment online. In line with this notion, Pauwels and Van Ewijk (2020) point out that information about consumer attitudes and intentions from survey responses remains highly relevant to managers. More specifically, an easy-to-apply parsimonious survey scale would enable managers to assess the relevance of different WOM types (including traditional offline WOM) in a uniform and standardized way, for various categories and consumer segments to manage WOM-related marketing activities and budget allocation. Likewise, such a scale would contribute to academic research by offering a way to examine the relevance of WOM to further deepen the field's understanding of the WOM phenomenon.

Against this background, this article makes three main contributions. First, from a conceptual perspective, we introduce the concept of "WOM relevance," which measures consumer perceptions of the relevance of WOM information for the purchase-decision process in a specific category. Thereby, we extend research on consumer perceptions of the relevance of certain marketing elements, such as brand (brand relevance, Fischer, Völckner, and Sattler 2010) and price (price consciousness, Ailawadi, Neslin, and Gedenk 2001). We develop and empirically validate a uniform instrument in the form of a parsimonious scale to measure WOM relevance at the individual consumer level. Our scale enables easy and comparable assessments of WOM relevance across consumers, categories, and WOM types, including newly emerging WOM types and those that are harder to capture (e.g., face-to-face WOM). We empirically show that WOM relevance consists of a category-level and a consumer-level component and that it comprises consumers' use of WOM information in the pre-purchase (i.e., search and consideration) and purchase (i.e., choice) phases, ultimately leading to the purchase of a product. By providing insights into consumer-level differences in WOM relevance, our scale can serve as a predictor, mediating, moderating, or outcome variable to advance the field's understanding of WOM-related consumer perceptions and behaviors. As a predictor variable, it can contribute, for example, to the literature on consumer information search by adding a construct that explains individual-level perceptions and behaviors regarding an omnipresent source of information. As a mediating variable, it can shed light on the underlying mechanisms that explain consumers' WOM-related behaviors. For example, it can help explain cross-category or cross-country differences in processing and using WOM information. As a moderating variable, it can help clarify when consumers rely on WOM and thus how effectively WOM influences attitudes and behaviors in the purchase-decision process for different consumers, categories, and WOM types. Finally, as an outcome variable, it can contribute to research on identifying consumer and category characteristics that facilitate WOM information search and usage.

Second, we empirically disentangle the consumer- and category-level components of WOM relevance, providing evidence that both components are important in consumers' purchase-decision process. We also show that the consumer-level variation is (much) larger than the category-level variation, which substantiates the importance of assessing WOM relevance at the consumer level—a key motivation for introducing the construct and developing the WOM relevance scale. The scale allows managers to target consumer segments with high or low WOM relevance differently even in categories wherein WOM relevance is generally high (or low).

Third, we illustrate that the scale can explain WOM-related behaviors, such as consumers' search behavior and consideration-set formation in an online-shopping environment, consumers' product choices, and consumers' intentions to retransmit others' WOM messages.

We develop, validate, and apply the WOM relevance scale in a series of eight studies (Table 1). Specifically, we use three studies from two countries (Studies 1–3) to develop and validate a parsimonious five-item scale. Next, Studies 4a and 4b disentangle the consumer- and category-level components of WOM relevance. While Study 4a explains the observed differences in WOM relevance across categories, consumers, and WOM types by a set of contingency factors, Study 4b examines the scale's relationship to related constructs and shows that these other constructs are unable to adequately capture consumer- versus category-level effects. Study 5 demonstrates the scale's predictive validity by showing that eWOM relevance relates to consumers' search behavior in an online-shopping environment. Consumers with high eWOM relevance rely more strongly on eWOM information by, for example, filtering or sorting based on products' star ratings. Likewise, eWOM relevance influences the composition of consumers' consideration sets resulting from their search behavior. Study 6 provides additional support for the scale's predictive validity by showing that it predicts hold-out choices as well as a more sophisticated choice model. Finally, Study 7 illustrates how the WOM relevance construct can advance the field's understanding of substantive WOM phenomena. Specifically, it shows that WOM relevance influences not only consumers' own purchase-decision process but also their intentions to retransmit others' WOM messages. Web Appendix A shows the sample characteristics for all studies and all items and the reliability estimates for each of the constructs.

Table 1. Overview of Studies

<table>
<tr><td>Studies 1–3</td><td>Scale Development
a) Item pool generation and item reduction
Literature review, roundtable discussions, and several iterations of item construction
Expert judgements (n = 10) to assess the content validity, comprehensibility, and completeness as well as the parsimony of item pool
Identification of seven non-redundant items based on expert judgements
b) Scale validation (Study 1)
German sample, n = 294, 10 categories
Confirmatory-factor analysis identifies five-item scale
c) Scale-revalidation: Additional language and categories (Studies 2 and 3)
U.S. representative sample, n = 414, 10 categories
German representative sample, n = 2275, 20 categories
Confirmatory-factor analysis reconfirms good model fit and high reliability
Demonstrates measurement invariance across categories and countries</td></tr>
<tr><td>Study 4a</td><td>Consumer- and Category-Level Components of WOM Relevance
German sample, n = 575, 20 categories, three per respondent
Disentangles the consumer-level and category-level components of WOM relevance: Both components significantly explain variation in WOM relevance, with the consumer-level component being (much) more important than the category component
Explains differences in WOM relevance by contingency factors
Shows that the nature of the relationship between WOM relevance and WOM availability is asymmetric</td></tr>
<tr><td>Study 4b</td><td>Comparing WOM Relevance with Related Constructs
US sample, n = 486, 20 categories, three per respondent
Disentangles the consumer-level and category-level components of WOM relevance
Shows that WOM relevance is different from related constructs (low correlations)</td></tr>
<tr><td>Study 5</td><td>The Role of WOM Relevance in the Search Stage
German representative sample, n = 307, one category
Predictive validity of WOM relevance scale for consumers’ search behavior
Demonstrates the relationship between WOM relevance and consumers’ search behavior and consideration-set formation in an online-shopping environment: eWOM relevance explains the use of search aids (e.g., using the customer review’s star rating to filter or sort products) and the composition of consideration sets (i.e., shopping baskets) in an online shop</td></tr>
<tr><td>Study 6</td><td>The Role of WOM Relevance in the Choice Stage
German sample, n = 2054, three categories
Predictive validity of WOM relevance scale for consumers’ purchase decisions
WOM relevance scale predicts product choices as well as a more sophisticated choice model</td></tr>
<tr><td>Study 7</td><td>The Role of WOM Relevance in Explaining WOM Retransmission
German sample, n = 440, one category
Investigating consumers’ retransmission intentions across WOM types
WOM relevance explains individual differences in WOM retransmission behavior for both the source and the target channel
WOM relevance influences not only consumers’ own purchase-decision process but also their intentions to retransmit others’ WOM messages</td></tr>
</table>

How Marketing Managers Think About WOM Relevance

While there is a vast amount of academic research on WOM, we first wanted to better understand how marketing managers think about the importance of WOM. More specifically, do they feel able to assess the relevance of WOM for consumers' purchase decisions and would they find a scientifically validated scale useful for management practice? How do they allocate budget to WOM marketing activities? We explored these questions in an online study among consultants from some of the largest consulting companies in Europe with experience in marketing and management. In total, 29 respondents with an average work experience of 12.7 years across a broad range of categories (e.g., automobile, retailing, groceries, FMCG, tourism, banking and insurance, fashion, and transportation) participated in the study. For more information about the sample and the questions asked, see Web Appendix B.

With a mean value of 4.24 and a relatively high standard deviation of 1.24 (7-point scale), the results indicate that many respondents are not convinced that managers can assess the relevance of different WOM types in a category. Respondents also do not consider measurement of the relevance of WOM in a specific category to be easy (mean value of 3.59, standard deviation of 1.80). Interestingly, when allocating budget to WOM activities in a specific category, managers seem to predominantly use the availability of WOM in that category as decision criterion (services: 5.41; durables: 5.68; non-durables: 5.86). Sales or profit generated in that category typically represent the second most frequently used criterion (services: 5.36; durables: 5.56; non-durables: 5.48). Even "gut-feeling" (services: 5.40; durables: 5.04; non-durables: 4.96) seems to be more important than actual information about the importance of WOM (services: 4.60; durables: 4.71; non-durables: 4.92). Accordingly, respondents mostly believe that a scientifically validated scale to measure the importance of different WOM types for consumers' purchase decisions in a product category would be helpful to effectively manage

WOM marketing activities (mean = 5.93; 1 = strongly disagree, 7 = strongly agree). They would also use the information provided by such a measurement instrument in their work (mean = 5.76; 1 = strongly disagree, 7 = strongly agree).

Overall, this explorative study suggests that marketing managers need an easy-to-apply measurement instrument that provides information about the relevance of different WOM types across categories, because they currently rely mainly on heuristics such as the availability of WOM or simple gut-feeling. Thus, it is the objective of this research to develop a parsimonious scale that enables managers to easily collect information about the relevance of WOM across consumers, categories and WOM types.

Conceptualizing WOM Relevance

Our conceptualization of WOM relevance focuses on the pre-purchase (i.e., search and consideration) and purchase (i.e., choice) phases, which eventually culminate in a consumer's purchase decision. When searching for product alternatives, evaluating them, and eventually choosing one, consumers are typically exposed to different pieces of information that vary in the extent to which they are viewed as relevant to consumers' purchase-decision process (e.g., Feldman and Lynch 1988). The perceived relevance of a piece of information in a consumer's purchase-decision process relates to the extent to which it helps the consumer discriminate between product alternatives (Herr, Kardes, and Kim 1991). Under the assumption that information from a specific WOM type provides a certain benefit to consumers (e.g., the reduction of perceived risk; Fischer, Völckner, and Sattler 2010), WOM relevance can be thought of as a general decision weight within the purchase-decision process that relates expected WOM benefits to other benefits, such as the benefit that results from a lower price. In line with this notion, we define WOM relevance as the extent to which the information from a certain WOM type (e.g., electronic WOM, or eWOM) influences consumers' decision making relative to other decision criteria in a specific category.

A key feature of this definition is that it captures both category- and consumer-specific components of the relevance of WOM information in a consumer's purchase-decision process. Namely, we expect that WOM relevance differs between categories depending on category characteristics, such as category type (e.g., durables, non-durables; e.g., You et al. 2015). At the same time, we expect consumers to differ in the extent to which they rely on WOM information in their purchase-decision process depending on their individual-level characteristics (e.g., Khare, Labrecque, and Asare 2011), such as opinion seeking. Thus, the consumer-level component can be interpreted as a trait-like consumer characteristic; it reflects the general relevance of WOM information in a consumer's purchase-decision process regardless of the category.

Finally, we expect that WOM relevance differs across WOM types. Drawing on established categorizations of communication channels used to spread WOM (e.g., Berger and Iyengar 2013; Hennig-Thurau, Wiertz, and Feldhaus 2015), we distinguish between three WOM types: (1) online customer reviews (i.e., eWOM), which include recommendations and product reviews published by consumers usually unknown to the reader on review or dealer websites, such as TripAdvisor and Amazon; (2) personal recommendations (i.e., traditional WOM, or tWOM), which include conversations with friends and colleagues held in person (i.e., face to face) or by telephone; and (3) recommendations from social media websites (i.e., social media WOM, or sWOM), which include comments and recommendations from friends, colleagues, and influencers ("friends" and "followers") in social networks, such as Instagram and Twitter. We believe these three generic types cover most of the WOM currently available for consumers and accordingly apply our scale to these three types. Note however, that, as our scale is WOM-type independent, it is easily adaptable to other (new) types of WOM.

Related Literature

There is broad research interest in WOM, and several authors have reviewed and synthesized the literature from different perspectives (e.g., Berger 2014; King, Racherla, and Bush 2014; Verma and Yadav 2021). Research on tWOM dates back to the pre-digital age with foundational studies, for example, by Arndt (1967); Herr, Kardes, and Kim (1991); and Richins (1983). With the rise of the internet, scholars started investigating eWOM, with a focus on consumer product reviews as well as messages in forums, blogs, and social media (Babić Rosario et al. 2020; Verma and Yadav 2021). Overall, the topics that WOM research covers can be divided into *three* major streams.

First, research on the senders of WOM has examined when and why consumers spread WOM and thus sheds light on *WOM availability*. For example, previous research has investigated motivational factors and reasons why consumers share WOM (e.g., Alexandrov, Lilly, and Babakus 2013; Hennig-Thurau et al. 2004), situational factors that influence people's sharing of WOM messages, such as product or service characteristics (e.g., Berger and Schwartz 2011), the sender's relationship with the receiver (e.g., Dubois, Bonezzi, and De Angelis 2016), and the channels via which consumers share WOM (e.g., Eisingerich et al. 2015), as well as the content that consumers share in different contexts (e.g., Ransbotham, Lurie, and Liu 2019).

Second, research on the *market impact of WOM* has quantified the effects of observable WOM characteristics (e.g., volume, valence, and variance) mostly from online channels on purchases and sales for different products and services (e.g., Liu 2006; Zhu and Zhang 2010). Meta-analyses on the sales impact of WOM indicate that the role of WOM may differ across category types (e.g., durable versus non-durable goods) and WOM types (e.g., consumer reviews versus social media posts). For example, You et al. (2015) find that WOM volume elasticities are higher for durables compared to non-durable, and Babić Rosario et al. (2016) find that WOM on e-commerce platforms has a stronger effect on sales than WOM on social media platforms.

Third, research on the *receivers of WOM* has investigated individuals' personal goals and motives to seek and use WOM (e.g., Hennig-Thurau et al. 2004), the impact of WOM on receivers' attitudes (e.g., Purnawirawan et al. 2015) and behaviors (e.g., Huang, Lurie, and Mitra 2009), and the underlying mechanisms that explain consumers' WOM-related behaviors (e.g., Laczniak, DeCarlo, and Ramaswami 2001). In addition, previous research has documented characteristics of senders (e.g., Weiss, Lurie, and MacInnis 2008), message content (e.g., Chen and Lurie 2013), and receivers (e.g., Khare et al. 2011; Guan and Lam 2019) that play a role in how effectively WOM influences consumers' attitudes and behaviors.

Scale-development endeavors within these three literature streams have focused on the senders of WOM (e.g., Goyette et al. 2010; Harrison-Walker 2001; Sweeney, Soutar, and Mazzarol 2012[1]). We are not aware of any systematically developed scale that addresses the receivers of WOM and, thus, can be used to assess the relevance of WOM in consumers' purchase-decision process. A few studies in the third stream use ad hoc scales to assess specific aspects of the influence of WOM on receivers. For example, Park and Lee (2009a) use three self-developed constructs to assess the perceived usefulness of eWOM, the frequency of eWOM use, and the stated effect of eWOM on final purchase decisions; Park and Lee (2009b) use the construct "eWOM effect" to assess the usage of a given eWOM message; and Khare et al. (2011) measure respondents' generic attitudes toward eWOM.

Our research adds to the second and third streams of literature. Specifically, while studies from the second stream demonstrate the sales impact of WOM and provide initial evidence that it differs across categories and WOM types, they cannot identify consumer-level differences in

[1] Sweeney et al. (2012) also consider the receiver perspective, but at the message-content level (similar to the sender perspective), which is not related to the relevance or importance of WOM.

WOM relevance. Our scale combines the category and consumer-level components of WOM relevance and thus help researchers and managers assess the role of both components in consumers' purchase-decision process. Furthermore, while studies from the third stream focus on assessing the impact of WOM in one category and for one specific WOM type, relatively few studies investigate, for example, product (e.g., Langan, Besharat, and Varki 2017; Sen and Lerman 2007; Zhu and Zhang 2010) or WOM-type differences (e.g., Baker, Donthu, and Kumar 2016; Yeap, Ignatius, and Ramayah 2014). However, none of these prior studies provide a uniform instrument to systematically quantify the relevance of WOM at the consumer level in different categories and for different WOM types.

Studies 1–3: Developing the WOM Relevance Scale

We develop the WOM relevance scale following established scale-development procedures (e.g., Churchill 1979; Netemeyer, Bearden, and Sharma 2003). Specifically, we aim to develop a reliable and valid scale that is parsimonious in terms of the number of items (e.g., Homburg, Schwemmle, and Kuehnl 2015) and is applicable to different WOM types (i.e., the three types considered in our study but also any other type that could potentially be relevant) and categories. The latter means that the scale should be generic in the sense that it is independent of the specific characteristics of particular WOM types and categories. In the following, we briefly summarize the procedure and results.

Based on a comprehensive literature review, roundtable discussions, and several iterations of item construction, we developed an initial item pool to represent the WOM relevance construct. As the aim was to develop a parsimonious scale that could be easily applied by managers and researchers alike, we needed to reduce the initial set of items. However, the typical scale purification using exploratory factor analysis and deletion of the items with the lowest loadings often leads to scales that do not capture the full domain of the focal construct but only include synonymous items of one sub-aspect. Therefore, an important step in the scale-purification process is "to ensure that the final set of items used captures or reflects the underlying construct as fully as possible, but without redundancy" (Lee and Hooley 2005, p. 368). Thus, we relied on two strategies to reduce the number of items: (1) expert judgements on content overlap (n = 10 academic experts) and (2) statistical purification using confirmatory factor analysis (CFA) using a sample of 294 qualified respondents recruited online in Germany. For more details see Web Appendix C.

Expert Judgements

We conducted in-depth interviews with academic experts to evaluate the content and face validity of our items. The experts read our definition of WOM relevance and then judged (1) the degree to which each of the items fit the conceptual definition of the WOM relevance construct, (2) the comprehensibility of the items, (3) the completeness of the item pool, and (4) the parsimony of the item pool (i.e., which items they deemed redundant). The experts judged most of our items as being appropriate according to the conceptual definition and found them comprehensible. Based on their feedback, we modified the wording of the items by applying a simplified and consistent sentence structure across all items, and we added two new items (Web Appendix Table C1). Regarding the parsimony of the item pool, the experts identified seven items as non-redundant facets of the construct's domain, which we test in the following Study 1.

Confirmatory-Factor Analysis (Study 1)

The results of the CFA confirm the factor structure with a single factor for each of the three WOM types (i.e., e/t/sWOM) for which the respondents evaluated the relevance within one randomly assigned category (out of a total of 10). However, the standardized loadings for Items 2 and 5 (Table C1) were much lower than those for the rest of the items across the three WOM types. As the objective was to develop a parsimonious scale, we deleted these two items, which improved model fit substantially ($\chi^2(72) = 108.3$, CFI = .990, RMSEA = .041, SRMR = .037). Table 2 shows the resulting five-item scale and Web Appendix Table C2 shows the internal consistency and reliability estimates for this scale.[2]

Finally, tests of discriminant validity showed that the new scale is distinct from potentially related constructs, such as WOM-type availability, WOM-type trustworthiness, brand relevance, category involvement, category expertise, and opinion seeking.[3] We also found discriminant validity among the three WOM types, providing evidence that they indeed represent different types of information that differ in relevance for consumers' purchase decisions (Table C3).

Table 2. Generic WOM Relevance Scale

No.	Introductory Text/Items
	The following statements are about the relevance of [WOM Type] for your purchase decision compared to all other pieces of information (e.g., [other WOM Types], advertising, sales talk/consultation, company website, packaging, etc.) that you may use for your decision. In this context, purchase decision refers to the entire process that ultimately leads to the purchase of a product—starting with the first interest in the category, continuing with the search for information and assessment of the range of available products up to the final decision for a particular product. If you think back to the last purchases of [Category], to what extent do you agree with the following statements? When I want to buy [Category], . . .
1	. . . [WOM Type] is very important to me.
2	. . . I search specifically for [WOM Type].
3	. . . [WOM Type] has a crucial influence on my purchase decision.
4	. . . [WOM Type] helps me better judge the quality of the product.
5	. . . [WOM Type] plays a significant role in my decision making.

Additional Language and Categories (Studies 2 and 3)

We revalidated the scale in a different language (Study 2, n = 414 qualified U.S. respondents) and on a more extensive set of 20 product categories (n = 2,275 qualified German respondents Study 3), using two representative samples provided by professional market research

[2] We focus on inter-item reliability. Future studies could additionally consider test-retest reliability.
[3] Table A2 provides an overview of all other items and the reliability estimates for each of the constructs in all studies.

companies. The results again indicate very good model fit as well as high internal consistency and reliability of the five-item scale (Web Appendix Table C2). Furthermore, we tested the scale's invariance with respect to different product categories.[4] The results show that the scale has metric and scalar invariance (Table C6), indicating the applicability of the scale to (potentially) many different product categories. Nevertheless, we find differences in means. While scalar and metric invariance are important for the comparability of the scale across categories, the differences in means are expected and in line with our assumption that WOM relevance differs across categories. Finally, although the aim of this research is not to compare WOM relevance across countries, we also examined the invariance of the scale across countries, which may be useful for future researchers applying the scale in such contexts. Overall, the results suggest that the scale exhibits metric and scalar invariance (Table C7).

Study 4a: Consumer- and Category-Level Components of WOM Relevance

In the following, we disentangle the consumer-level and category-level components of WOM relevance. In this study, respondents evaluated the relevance of e/t/sWOM in three randomly assigned categories (out of a total of 20) wherein they had recently purchased to provide insights into differences in WOM relevance across categories, consumers, and WOM types. We explain the observed differences by examining a set of six contingency factors.

Contingency Factor Framework

Based on theoretical models of external information search, we consider characteristics of the information source, the receiver, and the context as contingency factors. Specifically, prerequisites for using an information source are its availability or accessibility (Feldman and Lynch 1988) and, according to source credibility theory, its trustworthiness (Hovland and Weiss

[4] We could not perform this test in Study 1 due to a limited sample size per category.

1951). Regarding characteristics of the receiver, we consider category expertise, opinion seeking, and category involvement, which relate to consumers' ability and motivation to search for information (Schmidt and Spreng 1996). Finally, the information context (i.e., the product category) determines the perceived benefits and costs of search (Schmidt and Spreng 1996).

In terms of information source characteristics, we expect WOM availability to be one driver of WOM relevance because whether a piece of information eventually affects consumers' purchase-decision process should depend on its availability. However, availability does not necessarily imply relevance because not all available pieces of information are actually used in the decision process. Rather, a piece of information should only affect the purchase-decision process if it is perceived to be relevant for the decision at hand (Feldman and Lynch 1988). Likewise, consumers may perceive a piece of information to be relevant even if its availability is limited. For example, a consumer who perceives eWOM as highly relevant might extensively search for eWOM information even in categories wherein eWOM availability is limited. In addition, we consider the trustworthiness of the respective WOM type as a potential driver of WOM relevance. Trustworthiness influences a source's persuasiveness, and more trustworthy sources might play a greater role in consumers' purchase-decision process (You et al. 2015).

In terms of receiver characteristics, opinion seeking captures consumers' general tendency to seek advice from others when making a purchase decision (Flynn, Goldsmith, and Eastman 1996), so it might be positively related to WOM relevance. However, opinion seeking should not vary across categories. Rather, it refers to the subordinate trait-like importance of others' advice for consumers' decision making (Bearden, Netemeyer, and Teel 1989), whereas the relevance of WOM information for a consumer may vary across categories and WOM types. Category expertise refers to the level of knowledge that a consumer has about a certain product category (Nam, Wang, and Lee 2012). We expect a negative relationship between expertise and WOM relevance. If a consumer's category expertise is low (high), the relevance of WOM information for making a purchase decision in that category might be high (low). However, consumers with low category expertise might rely on other pieces of information (e.g., the brand) more strongly as well. Category involvement refers to consumers' interest in a specific product category (e.g., Laurent and Kapferer 1985) and thus impacts the amount of information consumers consider in their decision process. However, it does not capture the extent to which different pieces or sources of information eventually affect consumers' decision process in that category. Nevertheless, the two constructs might be related. For example, if a consumer's interest in a category is high, the relevance of WOM information in his or her purchase-decision process might be high as well. However, the relevance of WOM information could also be relatively low if other pieces of information (e.g., the brand) are viewed as more important.

Finally, for category type (i.e., context), we expect that WOM is more important for durables and services than for non-durables due to the higher purchase risk involved with these two category types, which makes consumers search for information (e.g., WOM information) more actively in these categories (Babić Rosario et al. 2016; You et al. 2015).

Model

We estimated the following basic model, where WOM-type relevance y_{ij} is composed of two group-level random intercepts for consumer *i* and category *j*. Equation 1 depicts a multi-level model (Gelman and Hill 2007; Goldstein 2011) that takes into account that there are repeated observations for each respondent and that each category was evaluated by several respondents. Since respondents were randomly assigned to three categories, the group-level error terms are non-nested independent effects, where t_{0i} is the consumer-level component and c_{0j} is the category-level component of WOM relevance, with $\mathrm{t_{0i}} \sim \mathrm{N}\left(0, \sigma_{\mathrm{t_0}}^2\right)$ and $\mathrm{c_{0j}} \sim \mathrm{N}\left(0, \sigma_{\mathrm{c_0}}^2\right)$.

$$\text{(e, t, s)WOM Relevance}_{ij} = y_{ij} = \alpha_{00} + t_{0i} + c_{0j} + \epsilon_{ij} \tag{1}$$

Equation 2 relates WOM relevance in category *j* by consumer *i*, y_{ij}, to the category-consumer-level, consumer-level, and category-level predictors. The category-consumer-level predictors (i.e., WOM-type availability, category expertise, and category involvement) are measured for consumer *i* and category *j*; they are category- and consumer-dependent drivers of WOM relevance. The consumer-level predictors (i.e., opinion seeking and WOM-type trustworthiness) are measured for consumer *i*; they are category-independent drivers that explain variation in the consumer-level component of WOM relevance (Equation 3). The category-level predictor (i.e., category type) is measured for category *j* and thus represents a consumer-independent driver of variation in the category component of WOM relevance (Equation 4).

$$y_{ij} = \alpha_{00} + \alpha_i^{consumer} + \alpha_j^{category} + \beta_1 \text{ WOM TYPE AVAILABILITY}_{ij} + \beta_2 \text{ EXPERTISE}_{ij} + \beta_3 \text{ INVOLVEMENT}_{ij} + \epsilon_{ij} \tag{2}$$

$$\alpha_i^{consumer} = \alpha_1 \text{ OPINION SEEKING}_i + \alpha_2 \text{ WOM TYPE TRUSTWORTHINESS}_i + t_{0i} \tag{3}$$

$$\alpha_j^{category} = \alpha_3 \text{ CATEGORY TYPE}_j + c_{0j} \tag{4}$$

In addition, Equation 5 considers a consumer-specific slope $\beta_{1i}^{consumer}$ that accounts for variation in the effect of WOM-type availability on WOM relevance across consumers. The specification of $\beta_{1i}^{consumer}$ (Equation 6) includes the consumer-level predictors (representing interactions of consumer characteristics and WOM-type availability) and an error term $t_{\beta i}$ that accounts for the unobserved variability across individuals. The consumer-level random intercept t_{0i} (Equation 3) and the consumer-level random slope $t_{\beta i}$ are allowed to correlate by ρ_t.

$$y_{ij} = \alpha_{00} + \alpha_i^{consumer} + \alpha_j^{category} + (\beta_{10} + \beta_{1i}^{consumer}) \cdot \text{WOM TYPE AVAILABILITY}_{ij} + \beta_2 \text{ EXPERTISE}_{ij} + \beta_3 \text{ INVOLVEMENT}_{ij} + \epsilon_{ij} \tag{5}$$

$$\beta_{1i}^{\mathrm{consumer}} = \beta_{11}\ \mathrm{OPINION\ SEEKING}_{i} + \beta_{12}\ \mathrm{WOM\ TYPE\ TRUSTWORTHINESS}_{i}\ +\ t_{\beta i} \qquad (6)$$

Data

We recruited 600 German-speaking respondents between the ages of 18 and 65 from Clickworker, a crowdworking platform similar to Amazon Mechanical Turk. We excluded respondents who self-reported a 2 or less on a scale measuring how much attention they paid to the survey (1 = almost no effort/attention, 5 = a lot of effort/my full attention), those who self-selected that their results should not be used in the data analysis, and respondents who clicked the same response category more than 80% of the time (i.e., straightliners [Paas, Dolnicar, and Karlsson 2018]), which left us with 575 qualified respondents.

All variables were entered into the model by computing the mean of the individual items except for category type, which was coded by five experts (see Web Appendix D for details on the expert coding). We grand mean-centered all consumer- and category-consumer-level variables and divided them by their standard deviation to obtain standardized estimates.

Results

We estimated the basic model described in Equation 1 (i.e., Model M1) to show that WOM relevance comprises a consumer-level and a category-level component. In addition, we estimated two variations of Model M1: one model that only includes the consumer-level component (M1a) and one model that only includes the category-level component (M1b). Furthermore, Model M2 depicted in Equations 2–4 includes the main effects of the consumer-level, category-level, and category-consumer-level predictors. Finally, Model M3 depicted in Equations 3–6 includes the main effects as well as the interaction effects between WOM-type availability and the two consumer-level predictors. We estimated all models using maximum likelihood with the R package lme4 (Bates et al. 2015). Table 3 depicts the results.

The results show that WOM relevance for all three WOM types is composed of a consumer-level and a category-level component. The AIC for the model that includes both the consumer-level and the category-level components (M1) improves model fit over each separate model (M1a, M1b). Likewise, a likelihood ratio test indicates that both components add significant explanatory power to the model. The standard deviations of the consumer- and category-specific error terms, σ_{t_0} and σ_{c_0}, show that the consumer-level variation (.600 – .770) is much higher than the category-level variation (.171 – .451) for all three WOM types. The strong contribution of the consumer-level component to the variation in WOM relevance substantiates the importance of assessing the role of WOM at the consumer level—a key motivation of developing the proposed WOM relevance scale. In addition, we found that the variation across categories is much higher for eWOM (.451) than for sWOM (.171) and tWOM (.211). Interestingly, sWOM shows the highest consumer-level variation (.770) of all three WOM types as well as the lowest category-level variation (.171).

Regarding the consumer-level predictors (i.e., Models M2 and M3), WOM-type trustworthiness has a significantly positive effect on WOM relevance for all three WOM types, whereas opinion seeking has a significantly positive effect on eWOM and tWOM relevance but is insignificant for sWOM relevance. Regarding the category-consumer-level predictors, WOM-type availability and category involvement have significantly positive effects on WOM relevance for all three WOM types. Interestingly, category expertise is significantly positively related to eWOM relevance but has no significant effect on tWOM and sWOM relevance. Expertise in a category may not substitute for WOM information. Rather, it seems to help consumers navigate through the large number of online reviews typically available for different products and assess their usefulness. In addition, we find (see Web Appendix D for details) that the relationship

Table 3. Main Effects and Interaction Effects of the Multi-Level Model (Study 4a)

	eWOM relevance					tWOM relevance					sWOM relevance				
	M1	M1a	M1b	M2	M3	M1	M1a	M1b	M2	M3	M1	M1a	M1b	M2	M3
Intercept	**3.215**	**3.147**	**3.217**	**2.700**	**2.691**	**3.246**	**3.219**	**3.244**	**3.044**	**3.055**	**2.226**	**2.199**	**2.247**	**2.060**	**2.069**
	(.106)	(.034)	(.103)	(.108)	(.108)	(.057)	(.032)	(.054)	(.059)	(.058)	(.053)	(.037)	(.063)	(.048)	(.046)
WOM-type availability				**.282**	**.289**				**.264**	**.250**				**.155**	**.176**
				(.024)	(.025)				(.022)	(.025)				(.021)	(.025)
Category expertise				**.069**	**.068**				-.003	-.005				.038	.043
				(.031)	(.031)				(.029)	(.029)				(.027)	(.027)
Category involvement				**.076**	**.071**				**.069**	**.065**				**.087**	**.082**
				(.032)	(.032)				(.030)	(.029)				(.027)	(.027)
WOM-type trustworthiness				**.299**	**.298**				**.180**	**.179**				**.492**	**.484**
				(.027)	(.027)				(.028)	(.028)				(.028)	(.028)
Opinion seeking				**.124**	**.123**				**.228**	**.238**				.013	.010
				(.027)	(.027)				(.028)	(.028)				(.027)	(.027)
WOM-type trustworthiness × WOM-type availability					**.064**					**.055**					.024
					(.022)					(.026)					(.023)
Opinion seeking × WOM-type availability					-.010					-.026					**.045**
					(.023)					(.024)					(.023)
Category type (durable)				**.708**	**.709**				**.162**	**.173**				**.203**	**.195**
				(.137)	(.137)				(.075)	(.074)				(.058)	(.055)
Category type (service)				**.646**	**.644**				**.387**	**.384**				**.249**	**.244**
				(.140)	(.140)				(.076)	(.074)				(.058)	(.055)
σ_{t_0}	.600	.593		.422	.413	.624	.628		.453	.432	.770	.791		.531	.515
σ_{c_0}	.451		.448	.223	.223	.211		.218	.101	.098	.171		.257	.066	.061
ϵ_{ij}	.827	.970	1.021	.777	.767	.747	.775	.973	.717	.676	.646	.665	.996	.634	.599
σ_{t_β}					.136					.275					.252
ρ_t					.498					-.174					.406
AIC	4778	5147	4952	4381	4373	4510	4575	4765	4174	4139	4314	4376	4848	3943	3900
Marginal R^2				.345	.345				.263	.262				.342	.344
Conditional R^2	.451	.272	.162	.525	.535	.438	.396	.048	.481	.538	.607	.594	.059	.616	.660

Notes: Standard errors are in parentheses; coefficients with $p < .05$ are marked in **bold**, marginal R^2 represents the variance explained by the fixed effects only, conditional R^2 is interpreted as the variance explained by the entire model, including both fixed and random effects between WOM relevance and availability is asymmetric for eWOM and sWOM, suggesting that availability is a necessary but not sufficient precondition for WOM relevance (e.g., Dul 2016). Regarding the category-level predictors, we found significantly positive effects for both durables and services compared to non-durables (reference category), as expected.

The results in Table 3 also reveal significantly positive interaction effects between WOM-type trustworthiness and WOM-type availability for eWOM and tWOM relevance. In other words, (e/tWOM) trustworthiness explains variation in the consumer-specific influence of (e/tWOM) availability, with higher trustworthiness resulting in a stronger effect of availability on WOM relevance. This interaction effect is plausible as consumers who trust a specific WOM type might vary in their WOM relevance due to the perceived availability of that WOM type. However, if consumers do not trust a specific WOM type, the extent to which information is available from this WOM type does not matter; the relevance of this (untrusted) WOM type is generally low. This may also explain the asymmetric nature of the relationship between availability and relevance found in our descriptive analysis.

Finally, Study 4a also provides interesting descriptive insights into WOM relevance at the category level (i.e., after partialling out the consumer-level variation). Table 4 shows the estimated average WOM relevance per category for each WOM type in descending order. We rescaled WOM relevance from the 1–5 scale to a 0–100 scale for better interpretability. While the spread confirms that eWOM relevance has larger variability than tWOM and sWOM relevance, the grand mean relevance scores for eWOM and tWOM are comparable, and they are higher than that for sWOM. In other words, eWOM relevance is much more differentiated across categories than tWOM relevance (or sWOM relevance, which is at a lower level). Furthermore, the rank order of categories reveals substantive differences between the relevance scores across WOM types. For example, consumers perceive tWOM to be highly relevant in the medical specialist category, while eWOM and sWOM are moderately relevant for this category (the category appears in the second half of the eWOM and sWOM rankings). In contrast, hotels rank first for eWOM but are in the second half of the ranking for tWOM.

Table 4. WOM Relevance Scores per Category and WOM Type (Study 4a)

Category	eWOM	Category	tWOM	Category	sWOM
Hotels	72.9	Medical specialists	67.6	Computer games	36.9
Small household appliances	66.7	Craftsmen	65.4	Movies (at the cinema)	36.5
Large household appliances	66.6	Movies (at the cinema)	61.7	Hotels	34.4
Mobile & smartphones	65.1	Insurance	60.3	Craftsmen	33.5
Computer games	63.8	Phone & internet providers	59.6	Mobile & smartphones	33.0
Insurance	61.6	Computer games	57.3	Designer sunglasses	32.1
Phone & internet providers	59.8	Large household appliances	57.1	Small household appliances	31.8
Leisure wear & shoes	58.9	Books	56.2	Books	31.7
Books	58.5	Cars	55.6	Cosmetics	30.9
Furniture	58.4	Small household appliances	55.5	Large household appliances	30.6
Long-distance transportation	57.5	Long-distance transportation	55.4	Phone & internet providers	30.6
Medical specialists	56.4	Hotels	54.8	Insurance	30.4
Cars	55.2	Furniture	54.6	Cars	30.4
Craftsmen	53.3	Mobile & smartphones	54.5	Furniture	30.1
Movies (at the cinema)	51.7	Cosmetics	54.1	Long-distance transportation	30.1
Designer sunglasses	50.9	Leisure wear & shoes	52.6	Leisure wear & shoes	29.9
Cosmetics	49.2	Designer sunglasses	51.4	Medical specialists	29.8
Body care	40.4	Body care	51.4	Body care	27.4
Detergent	35.1	Frozen food	49.8	Frozen food	22.2
Frozen food	25.8	Detergent	47.8	Detergent	20.6
Grand Mean	55.5	*Grand Mean*	56.1	*Grand Mean*	30.6
Spread ($Mean_{max} - Mean_{min}$)	47.1	*Spread ($Mean_{max} - Mean_{min}$)*	19.8	*Spread ($Mean_{max} - Mean_{min}$)*	16.2

Note: Values are rescaled from a 1–5 scale to a 0–100 scale for better interpretability.

Overall, these results show that WOM relevance comprises both consumer-level variation (a consumer-level component) and category-level variation (a category-level component). Furthermore, the consumer-level variation is generally (much) larger than the category-level variation. We can explain this variation using consumer-level variables, such as opinion seeking and WOM-type trustworthiness, as well as category characteristics, such as category type.

Study 4b: Comparing WOM Relevance with Related Constructs

In this study, we compare the WOM relevance scale to related constructs from the literature that capture certain aspects of WOM's impact on receivers (see the "Related Literature" section).

We compare the scale to respondents' generic attitudes toward eWOM (Khare et al. 2011) and Park and Lee's (2009a) three self-developed constructs that assess perceived usefulness of eWOM, frequency of eWOM use, and the stated effect of eWOM on the final purchase decision.

Study Design and Model

We use a similar design as in Study 4a. Respondents first answered the four related constructs from prior literature and then provided answers on eWOM relevance in three out of 20 randomly assigned categories wherein they had recently purchased. We focused on eWOM relevance to keep the questionnaire length reasonable and because (to the best of our knowledge) there are no similar s/tWOM constructs in the literature.

As in Study 4a, we disentangle the consumer-level and category-level components of WOM relevance (Equation 1). In addition, we estimate four models, in each of which we add one of the related constructs as a consumer-level predictor (Equation 7). We standardize β_1 using the consumer-level standard deviation $\sigma^2_{t_0}$ of the outcome. As we only have one predictor at the consumer-level (i.e., a univariate model) this standardized coefficients represents a correlation between the consumer-level component of WOM relevance and the predictor.[5]

$$\alpha_i^{\text{consumer}} = \beta_1 \text{ RELATED WOM CONSTRUCT}_i + t_{0i} \tag{7}$$

Data

We recruited 530 U.S. respondents from Prolific who indicated that English was their first language. A total of 514 respondents completed the survey. We excluded 28 respondents based on the same criteria as in Study 4a, which left us with 486 qualified respondents.

Results

We estimated the basic model described in Equation 1 (i.e., Model M1 and the two variations M1a and M1b) to show that eWOM relevance comprises a consumer-level and a

[5] We have run several simulations to ensure that we are able to retrieve the correct parameter.

category-level component (Table 5). The results are highly similar to Study 4a (σ_{t_0}= .644 in Study 4a and .600 in Study 4b; σ_{c_0}= .449 and .451). We also find a similar ranking of categories in terms of eWOM relevance with a high correlation of .85. While the U.S. sample in this study generally shows slightly higher eWOM relevance, there are noticeable difference for only three categories. In the U.S. sample (compared to the German sample in Study 4a), we find a substantially higher eWOM relevance for (designer) sunglasses and craftsmen, and a substantially lower eWOM relevance for insurance products.

As for the correlations between the WOM relevance scale and the related constructs, we find weak correlations at the overall sample level (.269 – .348; i.e., the correlation between y_{ij} and each alternative construct pooled across all categories J; see Web Appendix Table E1). This result was to be expected as we compare individual-level constructs (the related constructs) with eWOM relevance that exists at the category *and* consumer levels. When assessing the correlation between the consumer-level component of eWOM relevance and each of the related constructs based on the multi-level model results (Equation 7), we find higher correlations (.495 – .628; Models 2–5 in Table 5). These higher correlations suggest that the related constructs are associated with something like the consumer-level component of WOM relevance, but are not able to fully capture the entire concept. WOM relevance has the advantage of existing at both levels, thereby allowing to represent consumer-specific and category-specific effects.

Looking at the correlations within categories (i.e., the correlation between y_{ij} and each alternative construct for each category j; see Web Appendix Table E2), we find that the related constructs perform particularly poorly in capturing eWOM relevance in categories where WOM is less relevant or less prevalent in the market (e.g., body care, detergents, medical specialist). They align somewhat better with eWOM relevance in categories such as household appliances, computer games or movies, which are categories with a high prevalence of eWOM on e-commerce platforms such as Amazon or online rating portal such as Rotten Tomatoes or IMDB. Therefore, researchers and managers in categories outside of these "typical" WOM categories, in particular, should benefit from assessing the relevance of WOM using the proposed new scale rather than the related constructs. The latter seem to be based more on consumers' assessment of eWOM in a few highly salient categories that do not always match the purchase context at hand.

Table 5. Multi-Level Model Results (Study 4b)

	M1	M1a	M1b	M2	M3	M4	M5
Intercept	**3.357**	**3.359**	**3.355**	**3.357**	**3.356**	**3.356**	**3.357**
	(.107)	(.039)	(.104)	(.106)	(.105)	0.106)	(.104)
Generic eWOM attitude				**.495**			
				(.054)			
eWOM influence					**.591**		
					(.052)		
eWOM usefulness						**.540**	
						(.053)	
eWOM frequency							**.628**
							(.051)
σ_{t_0}	.644	.650		.560	.520	.542	.504
σ_{c_0}	.449		.449	.447	.444	.448	.440
ϵ_{ij}	.880	.983	1.091	.880	.880	.880	.880
AIC	4254	4458	4406	4177	4140	4161	4123
Marginal R^2				.073	.105	.087	.118
Conditional R^2	.444	.304	.145	.443	.442	.443	.441

Notes: Standard errors are in parentheses; coefficients with p < .05 are marked in **bold**

Study 5: The Role of WOM Relevance in the Search Stage

This study tests the predictive validity of the proposed WOM relevance scale in the search stage by showing that eWOM relevance relates to consumers' search behavior in an online-shopping environment. Specifically, the results show that consumers with high eWOM relevance have a higher likelihood of relying on eWOM information during search by, for example, filtering or sorting based on products' star ratings. Likewise, eWOM relevance influences the composition of consumers' consideration sets resulting from their search behavior.

Study Design

We implemented an online shop for vacuum cleaners that closely adapted design elements from the five largest online (consumer electronics) retailers in Germany. The online shop enabled users to search, browse, filter, and sort products (the default order was alphabetical by product name). To generate the corresponding product descriptions, we collected information from Amazon Germany on all vacuum cleaners in the €0–€200 price range, resulting in 451 products from 72 different brands (see Web Appendix Figure F1 for a screenshot of the online shop).

We implemented filter options for type (with or without bag), brand, price, review valence (average star rating), power (wattage), and allergy filter. We selected these six filter options because they are also implemented by at least three of the five major online retailers in Germany that cover 90% of the market share in online retail sales in consumer electronics. In addition, we implemented five sorting options (i.e., alphabetical, review valence, price upward, price downward, top seller [using Amazon's sales rank]) again because at least three of the five major online retailers in Germany offer these sorting options.

We instructed respondents that they should select five to 10 products into their consideration set (shopping basket) that they would seriously consider purchasing. Respondents could browse the shop, review their current shopping basket, and add (by clicking on "shopping basket" within the corresponding product snippet) or remove products as long as they wished. The selection of products into the consideration set was combined with an incentive-aligned mechanism adopted from Ding (2007). Specifically, we informed participants that we would raffle two €200 prizes at the end of the study. If they were among the winners of the €200, they would have to buy one randomly determined product from their consideration set. After reading some information about the study's objective (i.e., learning about their information behavior in the vacuum cleaner category) and the incentive-alignment procedure, respondents answered the first block of questions measuring either some consumer characteristics (e.g., online shopping experience, brand relevance; see Table A2 for the measures; Order Condition 1) or respondents' e/t/sWOM relevance (Order Condition 2). We randomized the order of the WOM relevance items and the consumer characteristics (which were roughly of same length) with one block appearing before the shopping task and one after to control for potential order effects. We do not find evidence for ordering effects (see Web Appendix F).

Data

A total of 313 respondents (307 qualified respondents; for exclusion criteria, see Study 4a) recruited from a German online access panel with an almost representative structure in terms of age and gender participated in this study. We collected clickstream data on every click a respondent made in the shop (e.g., selecting or de-selecting a filter or sorting option, selecting or de-selecting a product in the consideration set).

The prices of the products in respondents' consideration sets varied across the range of product prices in the shop. Importantly, we did not find a tendency toward lower prices, which may have indicated that respondents aimed to maximize their endowment payout (see the mean and median shop prices versus consideration-set prices in Web Appendix Table F1). Respondents seemed to balance price and product-performance characteristics to select high-quality products into their consideration sets. Overall, these findings (and additional analyses in Web Appendix F) indicate that respondents did not choose products randomly or with the intent to maximize their endowment payout but rather tried to select high-quality products. These findings also suggest that the incentive-aligned mechanism worked well.

Results

As expected, eWOM relevance relates to filter option and sorting selection. We ran logistic regression models to predict whether respondents selected review valence as a filter option (56%

of the sample) or sorted by review valence (15% of the sample) using e/t/sWOM relevance while controlling for brand relevance, price sensitivity, online shopping experience, and order condition (Table 6)[6]. We found a significantly positive effect for eWOM relevance (i.e., higher likelihood of selecting the review valence filter or sorting option) but not for tWOM or any of the control variables except for online shopping experience (i.e., higher likelihood of selecting the review valence filter with higher online shopping experience). In addition, we found a negative effect of sWOM relevance on selecting the review valence filter (but not on sorting by review valence). Thus, respondents with higher sWOM relevance were less likely to filter by review valence during search. Presumably, these people consider not only the content but also the sender of (online) WOM information, which conflicts with the anonymous nature of eWOM (i.e., personally unknown senders). As another plausibility check, we ran a logistic regression model to predict whether respondents selected a brand filter and found a significantly positive effect for the brand relevance construct (Fischer, Völckner, and Sattler 2010)—that is, higher brand relevance increases the likelihood of selecting a brand filter. Interestingly, higher eWOM relevance significantly decreases the likelihood of selecting a brand filter. Thus, respondents with higher eWOM relevance focused less on searching for specific brands.

Regarding the composition of respondents' consideration sets, respondents with higher eWOM relevance selected products with, on average, higher star ratings and smaller variation (and thus higher consistency) in the star ratings (i.e., lower standard deviation). Likewise, respondents with lower eWOM relevance selected products with higher variability and, on average, lower star ratings as they focused more on other attributes (e.g., price or brand). We ran

[6] The findings are robust to the inclusion of a Gaussian copula (Park and Gupta 2012) for eWOM relevance to address potential endogeneity concerns about unobserved confounders that could drive both WOM relevance and search behaviors. The results are available from the authors upon request.

regressions with the consideration set's mean star rating and its standard deviation as the dependent variables and e/t/sWOM relevance as predictors, again controlling for brand relevance, price sensitivity, online shopping experience, and order condition (Table 6). For the mean star rating, only the effect of eWOM relevance is significant. For the standard deviation, we found a significantly negative effect of eWOM relevance and a significantly positive effect of sWOM relevance, which is in line with the effect of sWOM relevance on filter usage described above.

Table 6. Results of (Logistic) Regressions

	Review valence filter	**Review valence sorting**	**Brand filter**	**Mean star rating consideration set**	**SD mean star rating consideration set**
Intercept	**-2.443** (.926)	**-5.504** (1.372)	**-2.109** (.925)	**3.769** (.152)	**.500** (.118)
eWOM relevance	**.779** (.149)	**.535** (.195)	**-.308** (.144)	**.128** (.023)	**-.077** (.018)
tWOM relevance	-.018 (.145)	.324 (.201)	.113 (.143)	.033 (.024)	-.013 (.018)
sWOM relevance	**-.286** (.134)	-.085 (.168)	.008 (.134)	-.034 (.023)	**.053** (.018)
Brand relevance	-.034 (.139)	.052 (.182)	**.854** (.149)	.042 (.023)	-.028 (.018)
Price sensitivity	-.106 (.169)	.090 (.232)	-.012 (.168)	-.055 (.028)	.042 (.022)
Online shopping experience	**.288** (.113)	.102 (.160)	.003 (.113)	.032 (.019)	-.003 (.014)
Order (condition 2)	.245 (.256)	-.061 (.330)	-.143 (.256)	.071 (.042)	.025 (.033)
R^2				.17	.11
Pseudo-R^2	.13	.07	.11		

Notes: SD = Standard deviation; Standard errors in brackets; significant effects on $p < .05$ in bold.

Overall, the results demonstrate that the WOM relevance scale is able to discriminate online search behavior. The use of product-review information (i.e., average star rating) in the search and consideration stage is strongly linked to eWOM relevance but is not linked or is even negatively linked to tWOM and sWOM relevance.

Study 6: The Role of WOM Relevance in the Choice Stage

To analyze the WOM relevance scale's correspondence with purchase decisions, we designed a discrete choice experiment in three categories to validate the scale against stated choices. In particular, we analyzed how well our scale can predict respondents' choices in a hold-out task and compared its performance against the more sophisticated discrete choice experiment.

Study Design

We considered three categories that cover the three category types from Study 4a (i.e., durables, non-durables, and services) and differ in WOM relevance (see Table 4): washing machines (representing durables from the large household appliance category), hotels (representing services), and deodorants (representing non-durables from the body care category). For each category, we collected data on e/t/sWOM relevance in the category using our WOM relevance scale as well as data from a choice experiment in which respondents completed 12 choice tasks, each consisting of three product alternatives. We used 10 choices to estimate a hierarchical Bayesian multinomial logit model to derive individual-level utility estimates for each attribute level. The other two choices were designed as hold-out tasks. To rule out ordering and demand effects, half of the respondents in the hotel and deodorant categories first completed the WOM relevance scale and then the choice experiment and vice versa for the other half. In the following, we report the results based on the combined dataset as the disaggregated results neither systematically affect respondents' WOM relevance perceptions nor the ability of the WOM relevance scale to predict respondents' choices in the hold-out task across the two ordering conditions and the aggregated dataset. We used a random design (with attribute overlap for washing machines and deodorants and without overlap for hotels)[7] to generate the product alternatives for each respondent. In the choice tasks, respondents were asked to select the product alternative they were most likely to buy.

The setting of the choice experiment was similar across the three categories. We used seven to eight attributes and up to five attribute levels to describe the product alternatives: four brand

[7] While a washing machine or deodorant brand likely offer several product alternatives with different attributes, this situation is unlikely for hotel brands (e.g., the same hotel brand with different star ratings in the same city).

names, five prices, three eWOM valences, three tWOM valences, three sWOM valences, and two to three additional attributes that were specific to the category (e.g., "spinning speed" and "energy efficiency" in the washing machine category). We carefully selected the attributes based on typical filter/search criteria used by online platforms, such as Amazon, as well as market research reports (statista.com). For the attribute levels, we used typical market representatives (i.e., important brands, price ranges around the market average, etc.). For a list of all attributes and attribute levels, see Web Appendix Table G1.

To assess how well the WOM relevance scale predicts respondents' choices compared to the more sophisticated discrete choice experiment, we used one of the hold-out tasks that only varied the e/t/sWOM attributes (Table G2). While the choice model derives utilities for all attribute levels, including the e/t/sWOM valence levels, the WOM relevance scale only provides information about consumers' general preferences for the three WOM types. Therefore, we used the hold-out task that kept the other product attributes constant to allow a fair comparison of the predictions based on the choice model and those based on the WOM relevance scale.

Method to Assess Predictive Validity

To make predictions based on the choice data, we estimated a hierarchical Bayesian multinomial logit model. Following the standard procedure, we used the first 10,000 draws as burn-in and the following 10,000 draws to estimate the posterior distribution, keeping every 10th draw. For each of the 1,000 posterior draws that we kept, we calculated each product's utility and corresponding choice probability for each respondent based on the first-choice rule as well as the logit rule (e.g., Green and Srinivasan 1990) using the values of the part-worth coefficients of that draw. We aggregated the choice probabilities across respondents to obtain the choice share of each product alternative for that draw. The obtained choice shares represent a posterior draw from the choice shares. With this posterior distribution of the choice shares, we were able to calculate the average choice share predictions and confidence intervals for these predictions using the 2.5% and 97.5% percentiles of the posterior distribution (Chapman and Feit 2015). To assess the predictive validity of the choice model, we calculated the mean absolute error (MAE) for each posterior draw, which is the deviation of each draw's choice share (logit or first choice) from the observed choice share in the hold-out task. Furthermore, we assessed the hit rate, which compares the first choices with the observed choices for all respondents and each posterior draw.

To make predictions based on the WOM relevance scale, we averaged the five items for each WOM type and respondent. A high WOM relevance score for a specific WOM type implies a preference for a product with positive WOM of that type, and this product should be preferred over a product with negative WOM of that type. We assume indifference for mixed WOM. For example, a product with positive tWOM, mixed sWOM, and negative eWOM (2.5 stars) should be chosen by respondents who have higher tWOM relevance than eWOM or sWOM relevance. Specifically, we applied the BTL rule (Green and Krieger 1988) to calculate choice probabilities:

$$\Pr(x_i) = \frac{\tilde{u}(x_i)}{\sum_{j \in S} \tilde{u}(x_j)} \tag{8}$$

The preference weight *u* for product x_i from the set of alternatives S is given by

$$u(x_i) = \sum_{t \in \{tWOM, sWOM, eWOM\}} v_t \times \Theta_{i,t} \tag{9}$$

where $\Theta_{i,t}$ represents the attribute level of product *i* (Equation 3), and v_t represents the WOM relevance score for each WOM type *t*.[8]

$$\Theta_{i,t} \begin{cases} 1, \text{ if WOM type } t \text{ of product alternative } i \text{ is positive} \\ 0, \text{ if WOM type } t \text{ of product alternative } i \text{ is mixed} \\ -1, \text{ if WOM type } t \text{ of product alternative } i \text{ is negative} \end{cases} \tag{10}$$

[8] A sensitivity analysis shows that in two of the three categories, a weight of 0 for mixed WOM has the highest predictive power (i.e., lowest MAE of the choice shares). We achieved a slightly better predictive power only in the deodorant category by setting the weight between .1 and .2. We also tested a value of 0 for both negative and mixed WOM, assuming that respondents would only consider positive WOM. The results are very similar.

As Equation 8 requires positive preference weights for each product alternative, we used $\tilde{u}(x_i)$, which is the normalization of $u(x_i)$, with $\tilde{u}(x_i) = u(x_i) - \min_{j \in S}\{u(x_j)\}$ so that the least preferred alternative gets a utility of 0 (e.g., Green and Krieger 1988). Otherwise, negative choice probabilities might occur.

We again used two approaches to predict the choice shares in the hold-out task based on the calculated choice probabilities. First, we aggregated the choice probabilities from the BTL model across respondents to derive choice shares (i.e., similar to the logit rule). Second, we applied the first-choice rule and assumed that respondents would choose the product for which they have the highest choice probability in the BTL model. If two alternatives had the same choice probability (i.e., same WOM relevance score), we assumed that respondents would randomly pick an alternative. Table G3 provides some calculation examples. We used the resulting choice-share predictions to calculate an average MAE and, in the case of the first-choice rule, also a hit rate.

Data

We collected data for the washing machine category using a German online access panel with an almost representative sample in terms of age and gender resulting in 371 qualified respondents out of 404 (for exclusion criteria, see Study 4a; we also excluded extreme speeders [i.e., less than half the median time] as it requires some time to carefully evaluate all alternatives and make reasonable choices). For the hotel (875 respondents, 807 qualified respondents) and deodorant (940 respondents, 876 qualified respondents) categories, we collected data using an online access panel from a professional market research company. Both samples are representative of the target population of Germany in terms of age and gender. We required respondents to have a recent purchase experience in the category (within the last three months for deodorants and the last year for hotels).

Results

Overall, we find very similar results across the three categories (Table 7). In the washing machine category, the predictive validity of the sophisticated choice model is satisfactory, with average MAEs of 4.1 (logit rule) and 3.0 (first-choice rule) and an average hit rate of 46.5% (significantly above the 33% benchmark for a random selection). However, the simple WOM relevance scale performs similarly well in predicting respondents' choices in the hold-out task, with the average MAE of the BTL model being slightly smaller (3.3) but within the confidence interval based on the choice-model predictions (Table 7, Panel A). In addition, the hit rate based on the predictions using the WOM relevance scale is even larger (58.4%) than the upper bound of the confidence interval of the choice model.

In the hotel category, the predictions using the WOM relevance scale perform equally well, with the average MAE of the BTL model being smaller than the lower bound of the MAE's confidence interval from the choice model (Table 7, Panel B). However, the hit rate is smaller (43.2%) than the hit rate from the choice model. A possible reason might be the appearance of ties between product alternatives when predicting choice shares based on the WOM relevance scale, which we resolved by averaging over the corresponding alternatives (i.e., assigning a 50% hit when two alternatives are on par and a 33% hit when all three alternatives are on par according to our assumption that respondents would randomly choose among equal alternatives). The choice model has no ties because the utilities always differ, indicating even small preferences for one or another WOM type, while our procedure of averaging over the alternatives drives the hit rate toward the naïve choice benchmark of 33%. The larger number of ties might also explain the lower hit rate in the hotel category (43.2%) compared to the washing machine category (58.4%), for which we observe fewer ties.

Table 7. Results of Predictive Validation (Study 6)

Panel A: Washing Machines						
		Predicted Choice Shares				Hit rate
		Product A	*Product B*	*Product C*	MAE	
CBC-HB	Logit model	39.4%	15.6%	45.0%	**4.1**	
		[35.5, 43.1]	*[13.0, 18.5]*	*[40.7, 48.9]*	*[2.3, 6.0]*	
	First choice	40.2%	12.9%	46.9%	**3.0**	**46.5%**
		[35.0, 45.3]	*[9.2, 16.7]*	*[41.2, 52.0]*	*[0.9, 5.8]*	*[42.6, 50.4]*
WOM Relevance	BTL model	46.7%	8.3%	45.0%	**3.3**	
	First choice*	50.1%	10.3%	39.6%	**6.0**	**58.4%**
Observed choice shares		**41.8%**	**9.7%**	**48.5%**		

* Ties were resolved by assigning a 50% hit in the case of two alternatives being on par (8.1%) and a 33% hit in the case of all three alternatives being on par (5.9%)

Panel B: Hotels						
		Predicted Choice Shares				
		Product A	*Product B*	*Product C*	MAE	Hit rate
CBC-HB	Logit model	37.7%	22.8%	39.5%	**3.7**	
		[35.5, 40.1]	*[20.6, 25.0]*	*[37.0, 41.9]*	*[2.3, 5.1]*	
	First choice	39.1%	19.9%	41.0%	**2.2**	**47.6%**
		[35.7, 42.5]	*[16.6, 23.3]*	*[37.4, 44.7]*	*[0.7, 4.1]*	*[45.1, 50.1]*
WOM Relevance	BTL model	41.5%	17.3%	41.2%	**0.7**	
	First choice*	42.3%	19.3%	38.4%	**2.6**	**43.2%**
Observed choice shares		**40.4%**	**17.3%**	**42.3%**		

* Ties were resolved by assigning a 50% hit in the case of two alternatives being on par (11.0%) and a 33% hit in the case of all three alternatives being on par (15.6%)

Panel C: Deodorants						
		Predicted Choice Shares				
		Product A	*Product B*	*Product C*	MAE	Hit rate
CBC-HB	Logit model	32.2%	27.7%	40.1%	**7.8**	
		[29.9, 34.3]	*[25.4, 30.0]*	*[37.9, 42.3]*	*[6.3, 9.3]*	
	First choice	31.8%	26.2%	42.0%	**6.8**	**41.3%**
		[28.5, 35.0]	*[22.7, 29.5]*	*[38.8, 45.2]*	*[4.5, 9.0]*	*[38.7, 44.0]*
WOM Relevance	BTL model	44.8%	16.9%	38.3%	**6.2**	
	First choice*	43.6%	20.5%	35.9%	**7.8**	**40.4%**
Observed choice shares		**36.4%**	**16.0%**	**47.6%**		

* Ties were resolved by assigning a 50% hit in the case of two alternatives being on par (8.3%) and a 33% hit in the case of all three alternatives being on par (24.2%)

Note: 95% confidence intervals in parentheses

In the deodorant category, the average MAE based on the BTL model is again smaller than the lower bound of the MAE's confidence interval based on the choice model (Table 7, Panel C). The hit rate is comparable to that from the choice model at 40.4%. In the deodorant category, we observe even more ties between the three product alternatives than in the hotel category. This result is mostly driven by the generally low WOM relevance in the deodorant category, which corresponds with a large number of respondents with 0 for their WOM relevance scores. In general, we find that it is more difficult to predict choices in the deodorant category than in the washing machine or hotel categories as WOM relevance in the former category is very low. Therefore, WOM relevance might not be a good criterion for choice decisions in this category.

Study 6 demonstrates that our scale predicts choices as well as a more sophisticated choice model. We therefore conclude that it has predictive validity for consumers' choice behavior.

Study 7: The Role of WOM Relevance in Explaining WOM Retransmission

The goal of Study 7 is to illustrate how the WOM relevance construct may help enhance the field's understanding of substantive WOM phenomena by serving as an important predictor and moderating variable. Specifically, we demonstrate the role of WOM relevance in explaining the retransmission of WOM information, which refers to passing on information about others' purchase and consumption experiences (e.g., Baker et al. 2016; De Angelis et al. 2012).

From a self-projection perspective, people who perceive a specific WOM type to be important likely assume that others perceive this WOM type to be important as well (Waytz and Mitchell 2011). We therefore expect that the perceived relevance of a specific WOM type influences consumers' retransmission intention in that WOM channel. For example, tWOM relevance should have a positive effect on consumers' tWOM retransmission intentions, and sWOM relevance should have a positive effect on consumers' sWOM retransmission intentions. Furthermore, we expect that the source channel (i.e., the channel from which the WOM information originates) moderates the effect of WOM relevance on retransmission intentions. Specifically, we expect a cross-channel WOM relevance effect such that sWOM (tWOM) relevance influences tWOM (sWOM) retransmission if the source channel is sWOM (tWOM). In other words, if there is a channel mismatch (i.e., the source channel is not the same as the retransmission channel), the extent to which consumers retransmit the corresponding WOM information also depends on the perceived WOM relevance of the source channel and not only on the WOM relevance of the retransmission channel. In addition, we argue that channel lock-in, known from consumers' information search and purchasing behavior (i.e., favorable attitudes toward searching for information on one channel translate into favorable attitudes toward purchasing on this channel [Verhoef, Neslin, and Vroomen 2007]), should also play a role in WOM retransmission behavior. That is, we expect that receiving WOM information from one channel translates into lower intentions to retransmit the information on another channel. In other words, we expect that channel mismatch (i.e., the source channel is not the same as the retransmission channel) translates into lower retransmission intentions.

Study Design

We randomly assigned participants to one condition in a 2 (WOM source channel: tWOM versus sWOM) × 2 (WOM valence: positive versus negative) between-subjects design. The context was a restaurant experience shared by a good friend. We manipulated the source channel and the valence of the restaurant experience using corresponding scenario descriptions. Specifically, in the tWOM (sWOM) condition, participants read that a good friend told them personally (wrote a social media post) about a nearby restaurant. In the positive (negative) valence condition, the restaurant experience was very good (very bad), and the friend is strongly considering revisiting (not revisiting) the restaurant (see Web Appendix Table H1 for the scenario descriptions). We measured respondents' intentions to talk about their friend's restaurant experience in a personal communication with other friends (tWOM retransmission) and their intentions to share the experience on social media (sWOM retransmission) on a seven-point scale (1 = "very unlikely," 7 = "very likely"). We do not include eWOM in this study, although it is a common source channel for restaurants (e.g., Yelp or Tripadvisor), because it is unlikely to be used as a retransmission channel (i.e., writing an online review about an experience of another person). Nevertheless, it would be interesting to study eWOM retransmission in future research.

As a covariate we assessed respondents' involvement with the restaurant category with three items (1 = "completely disagree," 7 = "completely agree"). As manipulation checks, respondents were asked to indicate the valence of the restaurant experience (1 = "very negative," 4 = "neutral," 7 = "very positive") and the source channel of the restaurant experience (i.e., personal communication, social media post, and three other irrelevant alternatives). Finally, respondents indicated their tWOM and sWOM relevance in the restaurant category.

Data

We recruited 500 German respondents from Clickworker between the ages of 18 and 65. Participants received monetary compensation according to the platform's statutes. The valence manipulation check shows that participants in the negative valence condition perceived the restaurant experience as strongly negative (M = 1.46, significantly lower than 4 = "neutral," $p < .001$), while participants in the positive valence condition perceived the restaurant experience as strongly positive (M = 6.27, significantly higher than 4 = "neutral," $p < .001$). The manipulation check for the source channel shows that 89.1% (89.4%) of the respondents in the tWOM condition (sWOM condition) identified the source channel correctly. As correctly processing the source channel is important for identifying the relevant WOM-type effects and interactions, we excluded the 53 respondents who did not pass the manipulation check of the source channel (in addition to excluding respondents who self-reported a 2 or less on a scale measuring how much attention they paid to the survey and those who self-selected that their results should not be used in the data analysis, see also Study 4a), leaving us with 440 qualified respondents.

Results

We ran regression analyses with sWOM and tWOM retransmission intentions as the dependent variables and a channel-mismatch dummy (which takes the value of 1 if the source channel is sWOM [tWOM] and the retransmission channel is tWOM [sWOM]), a valence dummy (negative = 1), and WOM relevance as predictors (Table 8). The first model (M1) includes the direct effects of the two dummy-coded experimental factors, their interaction effect, and the direct effects of sWOM and tWOM relevance. In the second model (M2), we added the interaction effects between channel mismatch and WOM relevance. Finally, the third model (M3) shows that the results are also robust to the inclusion of age, gender, and category involvement.

Table 8. Results of the Regression of Experimental Design Factors and WOM Relevance on Retransmission Intentions (Study 7)

	tWOM retransmission			**sWOM retransmission**		
	M1	**M2**	**M3**	**M1**	**M2**	**M3**
(Intercept)	**5.132**	**5.142**	**5.034**	**2.722**	**2.716**	**1.973**
	(.142)	(.141)	(.336)	(.138)	(.139)	(.325)
tWOM relevance	**.287**	**.279**	**.256**	.042	.103	.083
	(.064)	(.064)	(.065)	(.063)	(.089)	(.089)
sWOM relevance	**.155**	.044	.046	**.249**	**.245**	**.232**
	(.053)	(.073)	(.074)	(.051)	(.052)	(.053)
Channel mismatch	**-.679**	**-.705**	**-.684**	**-.550**	**-.534**	**-.577**
	(.200)	(.200)	(.198)	(.195)	(.196)	(.194)
Valence (negative)	-.235	-.298	-.307	**-.870**	**-.858**	**-.838**
	(.202)	(.203)	(.203)	(.196)	(.196)	(.193)
Channel mismatch × Valence (negative)	.122	.182	.210	**.563**	.534+	.514+
	(.284)	(.284)	(.283)	(.278)	(.279)	(.276)
Channel mismatch × tWOM relevance					-.116	-.104
					(.120)	(.119)
Channel mismatch × sWOM relevance		**.220**	**.197**			
		(.101)	(.100)			
Age			.011			**.014**
			(.006)			(.006)
Gender			-.201			.194
			(.141)			(.137)
Involvement			**.130**			**.170**
			(.054)			(.053)
R^2	.122	.132	.153	.122	.124	.157
Adj. R^2	.112	.120	.135	.112	.112	.139

Notes: Standard errors are in parentheses; coefficients with $p < .05$ are marked in **bold**; $^{+}$ $p < .10$.

As expected, the channel-mismatch dummy has a significantly negative effect on both tWOM (M1: b = –.679, $p < .001$) and sWOM (M1: b = –.550, $p = .005$) retransmission intentions.

In other words, retransmission intentions are significantly higher for the channel from which the respective WOM information originated, supporting our expectation about channel lock-in. We found neither an effect of valence nor an interaction effect between valence and channel mismatch on tWOM retransmission intentions. However, we find a negative direct effect of valence (M1: b = –.870, $p < .001$) and a positive interaction effect between valence and channel mismatch on sWOM retransmission intentions (M1: b = .563, $p = .043$; in Models M2 and M3, the interaction effect is marginally significant at $p = .057$ and $p = .060$). In other words, on social media, a different source channel (i.e., channel mismatch) reduces consumers' retransmission intentions only for positive experiences. For negative experiences, consumers show very low sWOM retransmission intentions regardless of the source channel (see Figure H1).

Most importantly, we find that WOM relevance influences consumers' intentions to retransmit the corresponding WOM type: tWOM relevance drives tWOM retransmission (M2: b = .279, $p < .001$), and sWOM relevance drives sWOM retransmission (M2: b = .245, $p < .001$). Thus, consumers who perceive tWOM (sWOM) as relevant to their own purchase-decision process are more likely to retransmit WOM information on this channel. In addition, we find a significant interaction effect between channel mismatch and sWOM relevance for consumers' tWOM retransmission intentions (M2: b = .220, $p = .029$). That is, sWOM relevance influences consumers' tWOM retransmission intentions if the source channel is sWOM. Thus, consumers who perceive sWOM as relevant to their own purchase-decision process are more likely to share WOM they received on social media with friends in personal communications. However, we do not find an interaction effect between channel mismatch and tWOM relevance on consumers' sWOM retransmission intentions. One reason might be that it is generally quite unlikely that someone would actively post another person's experiences on social media. Rather, people post about their own experiences on social media or share others' posts, but they rarely post about others' experiences they have heard about, whether or not they think tWOM is relevant.

In summary, Study 7 demonstrates that WOM relevance influences not only consumers' own purchase-decision process (as shown in Studies 5 and 6) but also their intentions to retransmit others' WOM messages and that WOM relevance can act as an important moderator to further differentiate consumers' WOM-related behaviors.

Discussion

Theoretical Contributions

This article introduces the concept of WOM relevance and develops, validates, and applies a parsimonious scale to measure WOM relevance at the consumer and category levels. In doing so, this research contributes to the literature on WOM by demonstrating that the relevance of WOM is specific to a particular WOM type and has two components: a consumer-level component, which represents the trait-like character of WOM relevance (i.e., the general relevance of each WOM type regardless of category), and a category-level component that resembles the relevance of a specific WOM type in a specific product or service category.

The finding that the consumer-level variation is (much) larger than the category-level variation substantiates the importance of a scale that assesses WOM relevance at the consumer level. Thereby, this research extends the literature on the market impact of WOM by showing that there is not only category-level variation in WOM relevance but also substantial consumer-level-variation, which studies at the aggregate market level cannot capture. However, the findings also underscore the importance of measuring WOM relevance at both the consumer and category levels. This need stems from the observation that the category-level variation in WOM relevance is still substantial (albeit smaller than the consumer-level variation). Constructs measured only at the consumer level are not able to capture this multi-level variation in WOM relevance.

Our conceptualization provides insights into differences in WOM relevance across consumers, categories, and WOM types as well as contingency factors that explain these differences. Thus, the findings extend work on contextual WOM factors by disentangling their influence at the consumer versus category level. This research also contributes to the literature on consumer information search by adding a construct that explains individual-level perceptions and behaviors regarding an omnipresent source of information. Finally, this research contributes to the emerging literature on the retransmission of WOM information by showing that WOM relevance explains differences in consumers' WOM retransmission behaviors.

Managerial Implications

Our scale-development and validation efforts provide managers with a reliable, valid, uniform, and parsimonious way to measure WOM relevance that is applicable to many different categories and WOM types, including those that are more difficult to track, such as traditional offline WOM. We demonstrate the applicability of the scale for a broad range of categories and three prevalent WOM types—eWOM, tWOM, and sWOM. However, as digital, social, and mobile media are rapidly evolving, other WOM types will likely emerge. As our scale is adaptable to any WOM type, it can accommodate these developments. Furthermore, with only five items, respondents can answer the scale quickly. Thus, managers can easily apply the scale to assess and track the relevance of different WOM types in a unified way in their categories. This information can help managers revise their marketing activities and create budget-allocation plans according to the role of WOM in relation to other purchase-decision criteria. Specifically, in categories wherein WOM is highly relevant, managers are well advised to pay special attention to managing WOM and provide enough resources for these activities relative to other decision criteria, such as traditional advertising, salesforce activities, and company website information. While our results show that the availability of WOM drives WOM relevance (Study 4a), we also find that the relationship is asymmetric and other consumer or category-specific factors influence WOM relevance. Therefore, using the availability (i.e., the volume of WOM) in a category may not be a good proxy for WOM relevance, which is what managers currently do.

Managing WOM is a highly complex task that requires managers to pay attention to both the consumer and category levels as well different WOM types and to monitor them separately. For example, our results show that within the hotel category, eWOM and tWOM are much more important than sWOM. Thus, marketing activities that, for example, provide satisfied customers with an incentive to post a review about their hotel stay on websites like Booking.com or TripAdvisor (i.e., eWOM) or that remind them to do so might be particularly effective. Likewise, promotional activities enhance the accessibility of a hotel stay in consumers' minds and, in turn, increase people's likelihood of talking about the respective hotel (Berger and Schwartz 2011). In contrast, information posted on Twitter or Instagram (i.e., sWOM) might be less effective (although the hotel category is relatively important within sWOM, sWOM relevance is generally lower than tWOM and eWOM relevance; see Table 4).

WOM relevance also differs across categories. For instance, non-durables have substantially lower WOM relevance scores than durables and services, as expected. Thus, managing WOM and providing enough resources for these management activities is more important for service providers and managers of durable products than for managers of non-durable products. Interestingly, sWOM is most relevant for services followed by durables (and non-durables). Although service providers like airlines and hotels typically respond quickly to user comments posted on company-owned social media channels (e.g., Facebook brand page), the relevance of sWOM implies that service providers should also monitor the web and respond to user comments posted on other (earned) social media channels, particularly if a user tags the company. As social media usage likely relates to sWOM relevance, managing sWOM in this way may create opportunities to positively influence consumers' engagement and, in turn, their economic activity with companies—something Manchanda, Packard, and Pattabhiramaiah (2015) show for online customer communities but may also generalize to other social media channels.

Furthermore, our finding that the consumer-level variation of WOM relevance is (much) larger than the category-level variation has important implications for market-segmentation decisions. For example, even if WOM relevance in the body care category is generally lower than WOM relevance in the large household appliances category, in both categories, consumers may differ in the extent to which they perceive WOM information as relevant for their individual decision making. Understanding the consumer-level (versus category-level) component of WOM relevance enables managers to identify consumer segments for which WOM information plays an important role even if WOM relevance in the category seems to be generally low. For example, when dealing with sWOM, the category-level variation is generally low, while the consumer-level variation is large, which may imply that managers should focus on the target consumers and not on the product category when managing sWOM. However, for managing eWOM, both the category and the target consumers are relevant. Relatedly, the consumer characteristics identified in this research might help managers better understand which types of consumers find WOM important and which types do not, which is not necessarily self-evident. For example, managers might intuitively think of high category expertise as a substitute for WOM information. However, our results show that category expertise is positively related to eWOM relevance (while it has no significant effects on tWOM and sWOM relevance). Category expertise could potentially help individuals navigate through online reviews and assess their trustworthiness.

Finally, our results show that eWOM relevance relates to consumers' search behavior and consideration-set composition in online-shopping environments, which has important implications for retailers. For example, retailers may want to implement easy-to-use sort and filter functions based on eWOM information, including not only average star ratings (which are typically available on most online-shopping websites) but also, for example, number of reviews or average usefulness of reviews. Furthermore, displaying average star ratings in the shopping cart (not only on the main product website) may help consumers refine their consideration sets.

Future Research

This article offers researchers a reliable, valid, uniform, and parsimonious scale that can be applied in future studies to further deepen our understanding of consumers' WOM-related perceptions and behaviors. For example, future research could apply the scale in additional countries, which would enhance our understanding of the influence of country-specific characteristics (e.g., economic variables, cultural values) on WOM relevance. Further, while our studies consider cross-sectional data in 20 categories, future studies could extend the application of the scale to additional categories to obtain an even more nuanced picture of WOM relevance. By implementing a longitudinal design, future studies might examine how the relevance of WOM information potentially evolves over time. Additionally, we call for research on other potential factors that influence the consumer-level component of WOM relevance and that could thus inform managers' market-segmentation decisions. Also, future research could use the scale to advance the field's understanding of tWOM versus online WOM. While the last two decades have mainly focused on understanding online WOM, less is known about the differences between online and offline WOM relevance for different consumers and categories. Finally, our scale measures the general importance of WOM information for purchase decisions in a specific category, but not for individual decisions between specific products. It thus also does not capture the relevance of positive and negative WOM separately. Future research could investigate whether there are differences in the relevance of positive and negative WOM to individual purchase decisions across consumers, categories, and WOM types.

References

Ailawadi, Kusum. L., Scott A. Neslin, and Karen Gedenk (2001), "Pursuing the Value-Conscious Consumer: Store Brands Versus National Brand Promotions," *Journal of Marketing*, 65 (1), 71–89.

Alexandrov, Aliosha, Bryan Lilly, and Emin Babakus (2013), "The Effects of Social-and Self-Motives on the Intentions to Share Positive and Negative Word of Mouth," *Journal of the Academy of Marketing Science*, 41 (5), 531–546.

Arndt, Johan (1967), "Role of Product-Related Conversations in the Diffusion of a New Product," *Journal of Marketing Research*, 4 (3), 291–295.

Babić Rosario, Ana, Kristine de Valck, and Francesca Sotgiu (2020), "Conceptualizing the Electronic Word-of-Mouth Process: What We Know and Need to Know about eWOM Creation, Exposure, and Evaluation," *Journal of the Academy of Marketing Science*, 48, 422–448.

Babić Rosario, Ana, Francesca Sotgiu, Krstine de Valck, and Tammo H.A. Bijmolt (2016), "The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors," *Journal of Marketing Research*, 53 (3), 297–318.

Baker, Andrew M., Naveen Donthu, and Vineet Kumar (2016), "Investigating How Word-of-Mouth Conversations about Brands Influence Purchase and Retransmission Intentions," *Journal of Marketing Research*, 53 (2), 225–239.

Bates, Douglas, Martin Mächler, Ben Bolker, and Steve Walker (2015), "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, 67 (1), 1–48.

Bearden, William O., Richard G. Netemeyer, and Jesse E. Teel (1989), "Measurement of Consumer Susceptibility to Interpersonal Influence," *Journal of Consumer Research*, 15, 473–481.

Berger, Jonah (2014), "Word of Mouth and Interpersonal Communication: A Review and Directions for Future Research," *Journal of Consumer Psychology*, 24 (4), 586–607.

Berger, Jonah and Raghuram Iyengar (2013), "Communication Channels and Word of Mouth: How the Medium Shapes the Message," *Journal of Consumer Research*, 40 (3), 567–579.

Berger, Jonah and Eric M. Schwartz (2011), "What Drives Immediate and Ongoing Word of Mouth?" *Journal of Marketing Research*, 48, 869–880.

Chapman, Chris and Elea McDonell Feit (2015), *R for marketing research and analytics*. Springer.

Chen, Zoey and Nicholas H. Lurie (2013), "Temporal Contiguity and Negativity Bias in the Impact of Online Word of Mouth," *Journal of Marketing Research*, 50 (4), 463–476.

Churchill Jr., Gilbert A. (1979), "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (1), 64–73.

De Angelis, Matteo, Andrea Bonezzi, Alessandro M. Peluso, Derek D. Rucker, and Michele Costabile (2012), "On Braggarts and Gossips: A Self- Enhancement Account of Word-of-Mouth Generation and Transmission," *Journal of Marketing Research*, 49 (4), 551–563.

Ding, Min (2007), "An Incentive-Aligned Mechanism for Conjoint-Analysis," *Journal of Marketing Research*, 44 (2), 214–223.

Dubois, David, Andrea Bonezzi, and Matteo De Angelis (2016), "Sharing with Friends Versus Strangers: How Interpersonal Closeness Influences Word-of-Mouth Valence," *Journal of Marketing Research*, 53 (5), 712–727.

Dul, Jan (2016), "Necessary Condition Analysis (NCA) Logic and Methodology of "Necessary but not Sufficient" Causality," *Organizational Research Methods*, 19 (1), 10-52.

Eisingerich, Andreas B., HaeEun Helen Chun, Yeyi Liu, He (Michael) Jia, and Simon J. Bell (2015), "Why Recommend a Brand Face-to-Face but Not on Facebook? How Word-of-Mouth on Online Social Sites Differs from Traditional Word-of-Mouth," *Journal of Consumer Psychology*, 25 (1), 120–128.

Feldman, Jack. M. and John G. Lynch (1988), "Self-Generated Validity and Other Effects of Measurement on Belief, Attitude, Intention, and Behavior," *Journal of Applied Psychology*, 73 (3), 421–435.

Fischer, Marc, Franziska Völckner, and Henrik Sattler (2010), "How Important Are Brands? A Cross-Category, Cross-Country Study," *Journal of Marketing Research*, 47 (5), 823–839.

Flynn, Leisa Reinecke, Ronald E. Goldsmith, and Jacqueline K. Eastman (1996), "Opinion Leaders and Opinion Seekers: Two New Measurement Scales," *Journal of the Academy of Marketing Science*, 24 (2), 137–147.

Gelman, Andrew and Jennifer Hill (2007). *Data analysis using regression and multilevel hierarchical models.* Cambridge University Press.

Goldstein, Harvey (2011), *Multilevel statistical models,* 4th ed. John Wiley & Sons.

Goyette, Isabelle, Line Ricard, Jasmin Bergeron, and François Marticotte (2010), "e-WOM Scale: Word-of-Mouth Measurement Scale for E-Services Context," *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, 27 (1), 5–23.

Green, Paul E. and Abba M. Krieger (1988), "Choice Rules and Sensitivity Analysis in Conjoint Simulators," *Journal of the Academy of Marketing Science*, 16 (2), 114–127.

Green, Paul E. and V. Srinivasan (1990), "Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice," *Journal of Marketing*, 54 (4), 3–19.

Guan, Chong and Shun Yin Lam (2019), "Product Rating Statistics as Consumer Search Aids," *Journal of Interactive Marketing*, 48 (1), 51-70.

Harrison-Walker, L. Jean (2001), "The Measurement of Word-of-Mouth Communication and an Investigation of Service Quality and Customer Commitment as Potential Antecedents," *Journal of Service Research*, 4 (1), 60–75.

Hennig-Thurau, Thorsten and Mark B. Houston (2019), *Entertainment Science*. Springer.

Hennig-Thurau, Thorsten, Caroline Wiertz, and Fabian Feldhaus (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 (3), 375–394.

Hennig-Thurau, Thorsten, Kevin P. Gwinner, Gianfranco Walsh, and Dwayne D. Gremler (2004), "Electronic Word-of-Mouth via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet?" *Journal of Interactive Marketing*, 18 (1), 38–52.

Herr, Paul M., Frank R. Kardes, and John Kim (1991), "Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective," *Journal of Consumer Research*, 17 (4), 454–462.

Hewett, Kelly, William Rand, Roland T. Rust, and Harald J. van Heerde (2016), "Brand Buzz in the Echoverse," *Journal of Marketing*, 80 (3), 1–24.

Homburg, Christian, Martin Schwemmle, and Christina Kuehnl (2015), "New Product Design: Concept, Measurement, and Consequences," *Journal of Marketing*, 79 (3), 41–56.

Houston, Marc B., Ann-Kristin Kupfer, Thorsten Hennig-Thurau, and Martin Spann (2018), "Pre-Release Consumer Buzz," *Journal of the Academy of Marketing Science*, 46 (2), 338–360.

Hovland, Carl I., and Walter Weiss (1951), "The Influence of Source Credibility on Communication Effectiveness," *Public Opinion Quarterly*, 15 (4), 635–650.

Huang, Peng, Nicholas H. Lurie, and Sabyasachi Mitra (2009), "Searching for Experience on the Web: An Empirical Examination of Consumer Behavior for Search and Experience Goods," *Journal of Marketing*, 73 (2), 55–69.

Khare, Adwait, Lauren I. Labrecque, and Anthony K. Asare (2011), "The Assimilative and Contrastive Effects of Word-of-Mouth Volume: An Experimental Examination of Online Consumer Ratings," *Journal of Retailing*, 87 (1), 111–126.

King, Robert A., Pradeep Racherla, and Victoria D. Bush (2014), "What We Know and Don't Know About Online Word-of-Mouth: A Review and Synthesis of the Literature," *Journal of Interactive Marketing*, 28 (3), 167–183.

Langan, Ryan, Ali Besharat, and Sajeev Varki (2017), "The Effect of Review Valence and Variance on Product Evaluations: An Examination of Intrinsic and Extrinsic Cues," *International Journal of Research in Marketing*, 34 (2), 414–429.

Laczniak, Russell N., Thomas E. DeCarlo, and Sridhar N. Ramaswami (2001), "Consumers' Responses to Negative Word-of-Mouth Communication: An Attribution Theory Perspective," *Journal of Consumer Psychology*, 11 (1), 57–73.

Laurent, Gilles and Jean-Noël Kapferer (1985), "Measuring Consumer Involvement Profiles," *Journal of Marketing Research*, 22 (1), 41–53.

Lee, Nick and Graham Hooley (2005), "The Evolution of "Classical Mythology" Within Marketing Measure Development," *European Journal of Marketing*, 39 (3/4), 365–385.

Liu, Yong (2006), "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing*, 70 (3), 74–89.

Manchanda, Puneet, Grant Packard, and Adithya Pattabhiramaiah (2015), "Social Dollars: The Economic Impact of Customer Participation in a Firm-Sponsored Online Customer Community," *Marketing Science*, 34 (3), 309–472.

Moliner-Velázquez, Beatriz, Maria Fuentes-Blasco, and Irene Gil-Saura (2021), "Segmenting Customers According to Online Word-of-Mouth about Hotels," *Service Business*, 15 (1), 103–130.

Nam, Myungwoo, Jing Wang, and Angela Y. Lee (2012), "The Difference between Differences: How Expertise Affects Diagnosticity of Attribute Alignability," *Journal of Consumer Research*, 39 (4), 736–750.

Netemeyer, Richard G., William O. Bearden, and Subhash Sharma (2003), *Scaling Procedures. Issues and Applications*. SAGE Publications.

Paas, Leonard J., Sara Dolnicar, and Logi Karlsson (2018), "Instructional Manipulation Checks: A Longitudinal Analysis with Implications for MTurk," *International Journal of Research in Marketing*, 35 (2), 258–269.

Park, Sungho and Sachin Gupta (2012), "Handling endogenous regressors by joint estimation using copulas," *Marketing Science*, 31 (4), 567–586.

Park, Cheol and Thae Min Lee (2009a), "Antecedents of Online Reviews' Usage and Purchase Influence: An Empirical Comparison of US and Korean Consumers," *Journal of Interactive Marketing*, 23 (4), 332–340.

Park, Cheol and Thae Min Lee (2009b), "Information Direction, Website Reputation and eWOM Effect: A Moderating Role of Product Type," *Journal of Business Research*, 62 (1), 61–67.

Pauwels, Koen and Bernadette Van Ewijk (2020), "Enduring Attitudes and Contextual Interest: When and Why Attitude Surveys Still Matter in the Online Consumer Decision Journey," *Journal of Interactive Marketing*, 52 (1), 20–34.

Purnawirawan, Nathalia, Martin Eisend, Patrick De Pelsmacker, and Nathalie Dens (2015), "A Meta-Analytic Investigation of the Role of Valence in Online Reviews," *Journal of Interactive Marketing*, 31 (1), 17–27.

Ransbotham, Sam, Nicholas H. Lurie, and Hongju Liu (2019), "Creation and Consumption of Mobile Word of Mouth: How are Mobile Reviews Different?" *Marketing Science*, 38 (5), 773–792.

Richins, Marsha L. (1983), "Negative Word-of-Mouth by Dissatisfied Consumers: A Pilot Study," *Journal of Marketing*, 47 (1), 68–78.

Schmidt, Jeffrey B. and Richard A. Spreng (1996), "A Proposed Model of External Consumer Information Search," *Journal of the Academy of Marketing Science*, 24 (3), 246–256.

Sweeney, Jillian. C., Geoffrey N. Soutar, and Tim Mazzarol (2012), "Word of Mouth: Measuring the Power of Individual Messages," *European Journal of Marketing*, 46 (1/2), 237–257.

Sen, Shahana and Dawn Lerman (2007), "Why are You Telling Me This? An Examination into Negative Consumer Reviews on the Web," *Journal of Interactive Marketing*, 21 (4), 76–94.

Todorov, Georgi (2021), "Word of Mouth Marketing: 49 Statistics to Help You Boost Your Bottom Line," (accessed December 20, 2022), https://www.semrush.com/blog/word-of-mouth-stats/.

Verhoef, Peter C., Scott A. Neslin, and Björn Vroomen (2007), "Multichannel Customer Management: Understanding the Research-Shopper Phenomenon," *International Journal of Research in Marketing*, 24 (2), 129–148.

Verma, Sanjeev and Neha Yadav (2021), "Past, Present, and Future of Electronic Word of Mouth (EWOM)," *Journal of Interactive Marketing,* 53 (1)*,* 111–128.

Waytz, Adam and Jason P. Mitchell (2011), "Two Mechanisms for Simulating Other Minds: Dissociations Between Mirroring and Self-Projection," *Current Directions in Psychological Science*, 20 (3), 197–200.

Weiss, Allen M., Nicholas H. Lurie, and Deborah J. MacInnis (2008), "Listening to Strangers: Whose Responses are Valuable, how Valuable are They, and Why?" *Journal of Marketing Research*, 45 (4), 425–436.

Westbrook, Robert A. (1987), "Product/Consumption-Based Affective Responses and Postpurchase Processes," *Journal of Marketing Research,* 24 (3), 258–270.

Yeap, Jasmine A. L., Joshua Ignatius, and T. Ramayah (2014), "Determining Consumers' Most Preferred eWOM Platform for Movie Reviews: A Fuzzy Analytic Hierarchy Process Approach," *Computers in Human Behavior*, 31, 250–258.

You, Ya., Gautham G. Vadakkepatt, Amit M. Joshi (2015), "A Meta-Analysis of Electronic Word-of-Mouth Elasticity," *Journal of Marketing*, 79 (2), 19–39.

Zhu, Feng and Xiaoquan (Michael) Zhang (2010), "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing*, 74 (2), 133–148.