When and Why Paid Reviews Are Bad Investments: the Impact of Monetary Incentives on Reviewer Certainty

Christiline du Plessis, Erasmus University Rotterdam, The Netherlands
David Dubois, INSEAD, France

Three experiments demonstrate that compared to monetary incentives (judged significant) or no incentives, small monetary incentives decrease review writers’ feelings of legitimacy, which results in greater expressed uncertainty in the reviews. In turn, this shift in content decreases recipients’ perception of the products reviewed and negatively affects their purchase intentions.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/1020025/volumes/v43/NA-43

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com.
How People Talk About Products in Word of Mouth
Chairs: Grant Packard, Wilfrid Laurier University, Canada
Jonah Berger, University of Pennsylvania, USA

Paper #1: Following the Blind: How Expertise and Endorsement Style Impact Word of Mouth Persuasion
Grant Packard, Wilfrid Laurier University, Canada
Jonah Berger, University of Pennsylvania, USA

Paper #2: When and Why Paid Reviews are Bad Investments: The Impact of Monetary Incentives on Reviewer Certainty
Christlinde du Plessis, Erasmus University, Netherlands
David Dubois, INSEAD, France

Paper #3: Write and Write Alike: How Personality Affects the Value of Word of Mouth
Zahra Tohidinia, University of Connecticut, USA
Nicholas Lurie, University of Connecticut, USA

Paper #4: Transmitting Well-Reasoned Word of Mouth Impairs Memory for Product Experiences
Zoey Chen, University of Miami, USA
Ann Schlosser, University of Washington, USA

SESSION OVERVIEW
What drives how people speak when they talk about products? And what are the consequences of the language consumers use in word of mouth settings? While it’s clear that word of mouth shapes consumer behavior (Chevalier and Mayzlin 2006; Chintagunta et al. 2013), and some work has begun to uncover behavioral drivers of word of mouth (Berger 2014), there has been less attention to the underlying role of language in these processes (cf. Moore 2012; Packard and Wooten 2013). This session integrates a variety of perspectives to shed light on how people talk about products in word of mouth and how these linguistic choices shape word of mouth’s influence on people.

Packard and Berger examine the language different types of consumers use to endorse products and how this can, paradoxically, lead word of mouth recipients to make worse choices. A combination of field data and laboratory experiments show that novices tend to use stronger, more persuasive, “I recommend” endorsements than expert consumers. This, in turn, leads word of mouth recipients to follow objectively inferior recommendations. While paying for online reviews might seem like a good idea, du Plessis and Dubois demonstrate that it actually undermines review persuasiveness. Two experiments show that the language used in paid (vs. unpaid) reviews produces heightened recipient perceptions that the sender is uncertain about their attitudes, increasing recipients’ own doubt of the product’s quality and decreasing their purchase intentions. A subsequent study demonstrates that incentive size moderates the effect and that paying senders a small (versus no or large) incentive makes them more uncertain of their attitudes because they feel less legitimate.

Tohidinia and Lurie leverage linguistic theory and text analysis to derive personality profiles based on the language people use in over 65,000 online reviews. This data and subsequent lab studies show, for example, that reviews containing language evincing ncuticism are particularly valued by word of mouth recipients, while those from linguistically agreeable reviewers are less so. Finally, Chen and Schlosser shed light on how language use in reviews impacts the reviewer. Using logical rather than vivid language impairs subsequent product memory. Subsequent studies demonstrate that this is due to the extent to which the different writing styles impact rehearsal of the reviewer’s original product experience.

Taken together, these papers offer insights highlighting the importance of new research examining how people talk about products. Given the increasing visibility of consumer language in online contexts, this session should appeal to a broad array of researchers, including those examining service interactions (e.g. email or live forum service), online brand communities, and social media. Beyond these directly relevant settings, ACR’s information processing contingent (e.g. advertising researchers) may wish to hear about new approaches to analyzing language use at the level of the word, clause, sentence or corpus. Finally, the papers may be of interest to those interested in discovering conceptual links between language and knowledge, certainty, memory, and personality.

We hope to engage the session audience in discussion about these papers and how researchers can examine consumer behavior through linguistic analysis. The following questions are offered to this end: How can researchers link prominent theories relevant to consumer behavior to linguistic analysis? How do communication modes (oral vs. written) impact language use in word of mouth? When does the linguistic sophistication of a speaker help or hurt interpersonal persuasion? In what ways does the interaction between speaker and audience states traits influence language use in persuasion settings?

Following the Blind: How Expertise and Endorsement Style Impact Word of Mouth Persuasion
EXTENDED ABSTRACT
Word of mouth is thought to help consumers make better choices (Berger 2014; Hennig-Thurau and Walsh 2003). Few companies will say their own products are terrible though, so relying on advertisements or company-generated communications isn’t always useful. Consequently, consumers turn to others’ opinions to help them select good options and avoid bad ones. But could the way consumers endorse products actually lead word of mouth recipients to make worse choices than they would have otherwise?

This paper examines the endorsement language consumers’ use, and the consequences such language has on word of mouth recipients. Sometimes people assert their positive product evaluation by saying “I like it” while other times they might say “I recommend it.” We suggest that these different “endorsement styles,” or implicit versus explicit expressions of product approval (McCrenken 1989), shape the persuasive impact of word of mouth. People should be more likely to choose someone else recommended rather than liked because recommendations are expected to come from authority and signal expertise (Brodsky et al 2010; Oxford 2015; Pornpitakpan 2004).

Importantly, however, we also suggest that consumer knowledge should influence the endorsement styles people use. Less knowledgeable consumers—that is, consumers with less category experience or expertise (Alba and Hutchinson 1987) should be particularly likely to explicitly “recommend” products to others because they are less aware that others’ tastes may differ from their own (i.e. preference heterogeneity; Feick & Higie 1992). Taken together, novices’ tendency toward more explicit endorsements, and social perceptions of explicitendorser’s heightened expertise, may lead word of mouth recipients astray. Five studies, including a mix of field data and laboratory experiments, test these predictions.
Study 1 analyzes nearly 1,000 real consumer online reviews to examine whether more experienced consumers are less likely to use explicit endorsements (e.g., “I recommend”). While more experienced consumers (+1SD in book purchase experience) explicitly recommended books only 6.5% of the time, this jumped to 35.7% among less experienced book consumers (-1SD; \( p < .001 \)).

Study 2 tests this relationship in the lab, and examines the hypothesized mechanism behind the effect. Participants in one condition were reminded that, “people have very different preferences when it comes to this kind of hotel” to prime preference heterogeneity. All participants were then asked to write a review of a hotel they had stayed at and enjoyed. Replicating Study 1, in the no prime condition, novices (-1SD in tested hotel expertise) were significantly more likely to “recommend” the hotel (\( M = 53.8\% \)) than experts (+1SD \( M = 31.2\% \); \( p < .05 \)). Reminding participants of preference heterogeneity, however, wiped out this difference (\( p = .36 \)).

Study 3 examines how explicit (“I recommend”) versus implicit (“I liked” or “I enjoyed”) endorsement styles impact persuasion, and the proposed mechanism (perceived expertise) underlying this effect. As predicted, a friend who provided an explicit (“I recommend it…”) restaurant endorsement was seen as more expert than an implicit (“I enjoyed it…”) endorser (\( p < .05 \)) and led people to believe they would be more likely to choose the restaurant (\( p < .001 \)). Bootstrap results confirm mediation of the impact of explicit endorsements on persuasion by perceptions of sender expertise (95% CI: .02, .17).

A final pair of studies uses a yoked design to test whether novices are more likely to use explicit recommendations, and whether when this is combined with their tendency to choose worse products, can lead to sub-optimal choices by word of mouth recipients.

In Study 4, we asked participants to choose between two wines. Both wines had three positive attributes, but to make one inferior we added three additional attributes pretested to be undesirable (e.g., “bottle contains clay sediment”). Experts should notice this and avoid the bad wine. However, low knowledge consumers use attributing numerosity as a cue for quality (Alba & Marmorstein 1987), and as a result should pick the inferior wine. Consistent with this expectation, novices (as measured by a wine knowledge test) were nearly twice as likely to pick the inferior wine as experts (55.3% vs. 30.3%; \( p < .01 \)).

Next, participants were asked to imagine that they enjoyed the wine they chose, and wanted to tell others about it. They indicated whether they would be more likely to say, “I recommend” or “I liked” the chosen wine. Novices were again more likely to use strong endorsements than experts (60.9% vs. 47.3%; \( p < .05 \)). In sum, wine novices both (1) picked an inferior product and (2) used a stronger endorsement style than experts to share their opinion about it.

Finally, in Study 5 we examined the impact of these endorsements on others’ choices. Participants were given the same wines as Study 4, but in addition to product information, half the participants (word of mouth condition) were also shown the aggregate percentage of consumers from Study 4 that “liked” or “recommended” as a summary of consumer opinion. As predicted, the word of mouth condition led to inferior choices. Compared to the product information only condition (41.9%), seeing how many others liked or recommended the wine led more people to choose the inferior wine (53.07%; \( p = .02 \)).

This research demonstrates that how people endorse products has important—and potentially aversive—consequences for word of mouth recipients. By looking at the mismatch between the drivers of endorsement style and their impact on recipients, this paper deepens our understanding of how and why language use impacts consumer attitudes and behaviors.

When and Why Paid Reviews are Bad Investments: The Impact of Monetary Incentives on Review Certainty

EXTENDED ABSTRACT

Many companies increasingly provide incentives to consumers for participating in advocacy programs or posting product reviews online. In a recent study, Amazon found that, of its 1,000 most active reviewers, 85% are incentivized and 92% of incentivized reviewers write reviews about the free products they received (Pinch & Kesler 2010).

How do incentives influence review effectiveness and effort? On the one hand, studies have found a significant positive correlation between incentive size and task performance (e.g., Heyman and Ariely 2004). Other works showed that people do not exert effort on a task unless it contributes to their own economic well-being (e.g., Lock et al 1980; Eisenhardt 1989; Gerhart et al 2009). A survey of top executives in different industries (\( N = 23 \)) confirmed that 70% of them believe that paying for reviews increases review effectiveness. Taken together, these findings suggest that monetary incentives may lead to higher quality reviews. On the other hand, researchers have argued that external rewards can be ineffective (see Pfeffer 1998) possibly because incentivized reviewers invest less effort in the review writing process (e.g., Lepper et al 1973; Deci et al 1999; Gneezy and Rustichini 2000; Vansteenkiste et al 2004; Gerhart and Fang 2015). These findings suggest that monetary incentives might decrease review effectiveness.

Beyond the role of effort, this research proposes that monetary incentives can alter the content of reviews by affecting how legitimate review writers feel when generating their review. Money is a social cue which signals reward in a similar manner to praise (Izuma et al 2008; Zink et al 2008; Saxe and Haushofer 2008). Furthermore, research has shown that the social meaning of money can be transferred to the self (Vohs, Mead and Goode 2006; Vohs, Mead and Goode 2008; Yang et al 2013). As such, word-of-mouth (WOM) senders may interpret monetary incentives as social signals. Specifically, we propose that the size of the incentive can influence senders’ feelings of legitimacy and result in shifts in their expressed uncertainty. In particular, we expect that low monetary incentives can trigger low feelings of legitimacy, compared to significant monetary incentives or no incentives. In turn, because feeling legitimate is a key antecedent to certainty—the sense of conviction that one’s attitudes are correct (Rucker et al 2014) – we hypothesize that low legitimacy will induce consumers to express greater uncertainty in their reviews. We propose that when people are paid, they rely on the reward to feel legitimate. A low reward might trigger consumers to lose their feelings of legitimacy, increasing expressed uncertainty. In contrast when consumers are not given an incentive (i.e., when social cues are absent) or when monetary incentives are significant (i.e., social cues are adequate), consumers might rely on their experience, resulting in greater certainty (Rucker et al 2014).

Three experiments tested the effect of monetary incentives on the certainty expressed in review content. Importantly, across experiments, WOM receivers were blind to whether senders had been paid or not, or how much they had been paid.

In study 1, participants (\( N = 86 \)) were either paid a bonus of $1 or no bonus to write a review about an online game. These reviews were later rated by 975 MTurkers (± 11 raters per review) on review positivity, objectivity and reviewer certainty. Random effects regression showed that unpaid WOM senders expressed more certainty in
the content of their reviews than paid WOM senders \( (p < .001) \). Importantly, the reviews did not differ in positivity or objectivity (all \( p \)-values > .05).

In study 2, 716 MTurkers were randomly presented one of the reviews from study 1 and rated how doubtful they were of the quality of the reviewed product, in addition to how uncertain they believed the reviewer was based only on the language used in the review. Random effects regression showed that people perceived paid senders to be more uncertain of their opinions than unpaid senders \( (p < .05) \). Importantly, paying the sender (vs. not paying) had a significant positive indirect effect on the receivers’ product quality doubts through perceptions of reviewer uncertainty (95% C.I. = .01, .18). In other words, paying senders to write a review increased receivers’ perceptions that senders were uncertain, based only on the language used in the review, which in turn made the receiver doubt the quality of the game.

In study 3 we test whether increasing the size of the incentive reverses the effect of incentive provision on attitude uncertainty. In addition, we also tested if legitimacy mediates the effect of providing a small incentive on attitude uncertainty. 205 Dutch students took this study in exchange for course credit and were randomly assigned to receive no payment, 10 cents or €10 as a bonus payment for writing a review. Participants were told to use a set of headphones to watch two videos, write a review of the headphones and then answer questions. We measured how uncertain people felt about the opinions they expressed in their review. Results revealed that paying a WOM sender a small (i.e., 10c) versus no or high (i.e., €0 or €10) incentive increased senders’ uncertainty \( (p < .05) \). Senders who received a small incentive (versus no incentive or a significant incentive) also felt that their opinion was less legitimate \( (p < .0001) \). Moreover, legitimacy mediated the effect of receiving a small incentive on attitude uncertainty (95% C.I. = [.03, .16]). Reviewers who were paid a small incentive were more uncertain of the opinions in their review than those who did not receive an incentive or received a large incentive, because the small incentive made them feel less legitimate.

Overall, the results contribute to the literatures on incentives, WOM, and certainty. First, we demonstrate that the quality of a review may stem from feelings of legitimacy and that providing an incentive shifts the source of legitimacy from one’s experience (internal) to the size of one’s reward (external). Second, we contribute to a growing body of work showing that uncertainty in WOM content can decrease consumers’ attitudes and purchase intentions (Dubois, Rucker and Tormala 2011). Third, we contribute to past work on consumers’ certainty by documenting for the first time how monetary incentives affects consumers’ certainty by affecting their feelings of legitimacy. Managerially we demonstrate how incentive size can affect the content of the review and its effectiveness.

**Write and Write Alike: How Personality Affects the Value of Word of Mouth**

**EXTENDED ABSTRACT**

Online word of mouth is considered a valuable source of information (Zhu and Zhang 2010); but not all word of mouth has an equally persuasive effect on consumers. Prior research shows that review characteristics; such as valence (Chevalier and Mayzlin 2006), length (Mudambi and Schuff 2010), product type (Forman et al. 2008; Zhu and Zhang 2010) and the presence of words indicating temporal contiguity between product experiences and review writing (Chen and Lurie 2013); affect the influence and perceived value of online reviews. However less has been done to explore whether and how reviewer characteristics, such as personality, affect the perceived value of online word-of-mouth.

Personality influences language content and word choice (Pennebaker and King 1999). Recent research shows that personality dimensions can be successfully extracted through automated text analysis based on words used in different long text collections such as blogs, emails, or stories (Tausczik and Pennebaker 2010; Cohn et al. 2004; Yarkoni 2010). However, there have been few attempts to extract personality from a single review. Prior research has shown that dimensions of personality can have positive as well as negative effects on message persuasiveness in face-to-face settings (Oreg and Sverdluk 2014); yet little is known about the persuasive role of personality in online word-of-mouth, a setting in which cues to personality beyond language use are often absent. In addition, there has been almost no effort to examine how the match between sender and receiver personalities affects persuasion. Using secondary data as well as lab experiments, we examine how reviewer personality, as expressed through language use, affects the perceived value of online word-of-mouth. We further examine how the interaction between writer and reader personalities influences review value.

In Study 1, we utilized Linguistic Inquiry Word Count (LIWC; Pennebaker et al. 2001) to analyze over 65,000 Yelp restaurant reviews. Personality scores for each review were calculated using weights from prior research on the relationship between LIWC word categories and the Big Five personality dimensions (Schwartz et al. 2013). The relationship between expressed personality scores and word-of-mouth value was assessed through a negative binomial distribution model in which value was measured by a count of the number of people who found the review useful. The results of this study show that extracted personality dimensions are significant predictors of review usefulness. Greater extraversion, conscientiousness, and openness have a significant positive relationship with review usefulness. Somewhat counter-intuitively, greater neuroticism also has a significant positive relationship with review value, while greater agreeableness has a significant negative relationship with review value.

Study 2 extended Study 1 to examine the effects of expressed personality on word of mouth value in a lab setting. In line with Study 1, openness and neuroticism had a positive effect on review usefulness while agreeableness had a negative effect on review usefulness. Further analysis shows that review valence, as proxied by the star rating (on a scale of 1-5, with 5 indicating a very positive experience) that accompanies the text of each review, mediates the relationships between neuroticism and review value as well as between agreeableness and review value. Greater neuroticism is associated with more negatively valenced reviews. Conversely, agreeableness is associated with more positive reviews. Consistent with prior research (Basuroy et al. 2003; Chen and Lurie 2013), negative reviews are perceived as more useful than positive reviews.

Study 3 examined whether compatibility between writer and reader personalities impacts review usefulness. This study also examined whether the extracted personality dimensions of writers are consistent with those perceived by readers. Dimension of writer personality were extracted from online reviews. Dimensions of reader personality were extracted by asking readers to compose either a short email or a short review—we found similar results regardless of writing type—and a reader-writer personality similarity index was calculated using the mean of absolute differences between reader and writer on the Big Five (Setterlund and Niedenthal 1993). Readers also assessed reviewer personalities using the 10-item Big Five scale (Rammstedt and John 2007). Results show that greater similarity between expressed writer and reader personalities positively
impacts the perceived value of online word-of-mouth. Results also show that, except for openness, reader perceptions of reviewer personality are consistent with those extracted from reviews.

This paper shows that personality, extracted from a single online review through automated text analysis, is an important predictor of the value of online word-of-mouth. In addition to removing the need to measure personality through traditional scales, this approach is less subject to exaggeration and social desirability biases (Bearden et al. 1993). Our findings show that personality dimensions can influence word-of-mouth persuasiveness in unexpected ways. Although neuroticism is generally perceived as an undesirable personal attribute that reduces persuasion (Costa and McCrae 1980; Oreg and Sverdlik 2014), in word-of-mouth contexts it enhances value by being associated with more negative reviews. Conversely, conventional thinking holds agreeableness to be a favorable personal trait that enhances persuasion (McCrae and Costa 1987; Oreg and Sverdlik 2014); yet it decreases review usefulness by being associated with more positive reviews.

**Transmitting Well-Reasoned Word of Mouth Impairs Memory for Product Experiences**

**EXTENDED ABSTRACT**

A growing stream of research examines how word of mouth (WOM) affects consumers and product performance (e.g. Chevalier & Mayzlin 2006). The influence of engaging in WOM on the speaker, however, has been largely ignored. Although some recent research has examined how transmitting WOM affects senders' attitudes (Moore 2012), little is known about how transmitting WOM affects senders' memories for product experiences. Research in eyewitness testimony (Lofstus 2003) and verbalization (Schoeller and Englster-Schoeller 1990) suggest that the act of talking about an experience can change people's memory of the experience. Yet, it is unclear how and why different verbalization motivations affect memory. This work thus attempts to both contribute to a better understanding of how WOM affects reviewers and also provides insights into understanding how and why verbalization affects memory.

Reviewers often aim to be logical when writing reviews. Indeed, this behavior is often encouraged by many review websites (e.g. Amazon, Yelp, Yahoo). For example, Yahoo Movie prompts reviewers for rationales: “Tell us if you liked or disliked the film, but also why you liked or disliked it. Compare this film to others - why is this better or worse?”

We theorize that writing reviews can affect memory by influencing how people rehearse the original product experience. When reviewers write reviews in general, especially when prompted to write logical reviews, they search for well-reasoned arguments for their opinions rather than rehearse the original product experience (which is encouraged by imagery-based instructions). As a result of this search, those who write logical reviews are likely to suffer memory issues. Removing reviewers' ability to engage in this reason-seeking process — via cognitive load — should then improve memory to a level comparable to those writing imagery-based reviews. A series of five studies test and confirm these ideas.

In study 1, all participants watched the movie “Oktapodi,” a short (2 minute) animation, and then were randomly assigned to a sender or receiver condition. Senders were instructed to write a review about the video to another student. Each receiver was yoked to one sender and read a randomly-assigned review after viewing the film. This provides a delay between watching the film and the memory test, and allows us to assess whether it is the act of writing a review or the actual content of the review that affects memory. All participants completed a recognition task that had seven true statements and seven false statements about the movie. Memory is measured by summing correct acceptances and correct rejections. Results show that reviewers had worse memory for the underlying film than those who did not write reviews (e.g. receivers). Study 1 provides initial evidence that writing reviews can impair memory.

In the subsequent studies, to test the underlying process, we focus on senders and directly manipulate the goal to write logic-based versus imagery-based reviews. In study 2, after watching Oktapodi, participants were randomly assigned to write a logic-based review (i.e., to be well-reasoned and logical) or imagery-based review (i.e., to visualize and describe the film). Participants then completed the same memory test used in study 1. Manipulation checks supported the effectiveness of the manipulation. Confirming our theory, those who wrote logic-based reviews had worse memory than those who wrote imagery-based reviews. These results are consistent with the idea that searching for rational arguments may be responsible for subsequent memory deficits.

Study 3 further documents the negative effect of writing logical reviews and shows that the magnitude of this impairment is comparable to not being able to consolidate the product memory in the first place. Participants were randomly assigned to one of three between-subject conditions (logic-based review, imagery-based review vs. interference task). The procedures were the same as study 2 except that instead of writing a review, those in the interference condition completed an unrelated (anagram) task that prevented them from further thinking about the film. All participants then completed a 14-item memory test. Results reveal that those who wrote logic-based reviews had worse memory than those who wrote imagery-based reviews. Importantly, those who wrote logic-based reviews did as poorly on the memory test as those who experienced memory interference. This study not only reinforces earlier results but provides insights into the magnitude of the deficit: writing logic-based reviews is akin to preventing consolidation of the original memory.

In study 4, we test people's memory for a camera to examine the robustness of our effect in a different product category (i.e., a search rather than experience good). As before, participants were first shown a video of a camera, asked to write a logic- or imagery-based review for the camera and then answered 14 true/false statements. The results replicate earlier findings and show that writing logic-based reviews impairs memory for search products as well.

If memory impairments are indeed caused by people searching for rational arguments (instead of reflecting on the original experience), then preventing people from this search should attenuate our effect. Study 5 tests this with a 2 (logic- vs. imagery-based review) x 2 (no vs. high cognitive load) between-subjects design. The procedures were the same as study 4 except that those in the high cognitive load condition were asked to remember a roster of names while writing the review (after watching the movie; Elder & Krishna, 2010). Immediately after writing the review, the load was removed. Results show that when there is no load, the results replicate earlier findings. For those under cognitive load (i.e., prevented from searching for argument) while writing their reviews, there was no difference across logic- and imagery-based conditions.

In summary, we show that writing reviews—logical reviews in particular—can make people more susceptible to memory issues for the reviewed product. This work provides theoretical insight into how and why verbalization affects memory, as well as managerial insights given that many websites encourage consumers to write logic-based reviews.
REFERENCES
Pinch, Trevor and Filip Kesler (2010), “How Aunt Ammy Gets Her Free Lunch: A Study of the Top-Thousand Customer Reviewers at Amazon.com,” https://docs.google.com/viewer?a=v&pid=sites&srcid=ZnJlZWx1bmNoLm1lfHRle3R8Z3g6NDg0MjliMjM3OWQzMWRI


Yarkoni, Tal (2010), “Personality in 100,000 Words: A Large-Scale Analysis of Personality and Word use among Bloggers,” Journal of Research in Personality, 44 (3), 363-73.

