The Effect of Hyperbole on the Perceived Helpfulness of Amazon Product Reviews

Katherine M. Pierce

Follow this and additional works at: https://digitalcommons.memphis.edu/etd

Recommended Citation
Pierce, Katherine M., "The Effect of Hyperbole on the Perceived Helpfulness of Amazon Product Reviews" (2024). Electronic Theses and Dissertations. 3461.
https://digitalcommons.memphis.edu/etd/3461

This Thesis is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khggerty@memphis.edu.
THE EFFECT OF HYPERBOLE ON THE PERCEIVED HELPFULNESS OF
AMAZON PRODUCT REVIEWS

by

Katherine M. Pierce, B.A.

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

Major: Psychology

The University of Memphis
May 2024
Acknowledgments

First, I sincerely thank my mentors, Dr. Gina Caucci and Dr. Roger Kreuz, to whom I owe my growth and success as a graduate student. Thank you both for pushing me out of my comfort zone and for your unwavering support, patience, and guidance. I am incredibly grateful for the role you both played as not only a mentor but as a counselor, confidant, and friend.

Specifically, I want to thank Dr. Caucci for her incredible dedication and perseverance in navigating the analyses portion of this project. She went above and beyond, helping and teaching me every step of the way, and I can confidently say that this project would not have come to completion without her. I also want to thank Dr. Leah Windsor for setting aside time to be a part of my thesis committee, believing in this project, and for her enthusiasm, encouragement, and insightfulness.

To my parents, John and Kendall Pierce – thank you both for your immeasurable support and encouragement during this process, even with 432 miles between us. My ability to succeed is only made possible by the sacrifices you both have made. I admire and love you both so much!

To Michael – thank you for always meeting me with unwavering patience and understanding, picking up slack when needed, and never letting me take life too seriously. You created the space for me to give this project my all and stood by me every step of the way. I couldn’t have made it to the finish line without you – I love you!

Finally, to my dog, Meg, who kept me company during the long days and even longer nights. She is all the best things in life wrapped into one, and I could not have made it without her by my side (or lying at my feet).

The last two years have been a clear demonstration of The Lord’s faithfulness. I will never be able to fully express my gratitude for each person who has helped me along the way.
Abstract

Hyperbole is frequently used in interpersonal communication, yet little is known about its communicative function in casual discourse. This study aims to clarify the function of hyperbole in everyday speech, particularly in persuasive communication. Using Amazon product reviews, this study explores the effect of hyperbole and the moderating effects of term valence (i.e., positive and negative hyperbole) and product type (i.e., search and experience goods), on review helpfulness, which serves the measure of persuasion. The presence of positive and negative hyperbolic terms in over 22,000 product reviews was computed via LIWC. Individual Tobit regression analyses were performed for hyperbole and moderating factors. Results indicate overall and positive hyperbole increases review helpfulness, while negative hyperbole decreases helpfulness. Further, negative hyperbole was found to decrease review helpfulness for search goods, whereas overall and positive hyperbole increased helpfulness for experience goods.
Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of Tables</td>
<td>V</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Hyperbole</td>
<td>4</td>
</tr>
<tr>
<td>Online Product Reviews</td>
<td>9</td>
</tr>
<tr>
<td>Determinants of Review Helpfulness</td>
<td>12</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>15</td>
</tr>
<tr>
<td>Methods</td>
<td>16</td>
</tr>
<tr>
<td>Term Valence Categorization</td>
<td>17</td>
</tr>
<tr>
<td>Dataset Description and Preprocessing</td>
<td>18</td>
</tr>
<tr>
<td>LIWC</td>
<td>21</td>
</tr>
<tr>
<td>Design</td>
<td>22</td>
</tr>
<tr>
<td>Results</td>
<td>25</td>
</tr>
<tr>
<td>The Effect of Hyperbole on Review Helpfulness</td>
<td>25</td>
</tr>
<tr>
<td>The Moderating Effect of Term Valence</td>
<td>25</td>
</tr>
<tr>
<td>The Moderating Effect of Product Type</td>
<td>27</td>
</tr>
<tr>
<td>Correlational Analysis: Hyperbolic Term Valence and LIWC Emotion Scores</td>
<td>29</td>
</tr>
<tr>
<td>Discussion</td>
<td>30</td>
</tr>
<tr>
<td>Implications</td>
<td>32</td>
</tr>
<tr>
<td>Limitations</td>
<td>32</td>
</tr>
<tr>
<td>Future Directions</td>
<td>33</td>
</tr>
<tr>
<td>References</td>
<td>35</td>
</tr>
<tr>
<td>Appendix</td>
<td>48</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Description of Datasets after Selection and Preprocessing</td>
<td>20</td>
</tr>
<tr>
<td>2. Variable Description and Operationalization</td>
<td>24</td>
</tr>
<tr>
<td>3. Tobit Regression Output: Effect of Overall, Positive, and Negative Hyperbole on Perceived Helpfulness</td>
<td>26</td>
</tr>
<tr>
<td>4. Tobit Regression Output: Effect of Product Type on the Relationship Between Hyperbole and Perceived Helpfulness</td>
<td>28</td>
</tr>
<tr>
<td>5. Correlation Between Positive and Negative Hyperbole and LIWC Positive and Negative Emotion Scores</td>
<td>30</td>
</tr>
</tbody>
</table>
Introduction

Interpersonal communication is fundamental both on a societal and individual level (Fussell & Kreuz, 1998; Ramaraju, 2012). Whether computer-mediated or face-to-face, the exchange of information between individuals is imperative for societal cohesion and personal growth (Gamble & Gamble, 2014). Discourse processing and pragmatics have emerged as focal points in the research literature, revealing the intricate nuances of interpersonal communication. Consequently, many researchers have turned their focus to figurative language, parsing apart the different figures of speech, such as metaphor, irony, and hyperbole, to better understand the discourse goals they accomplish.

Hyperbole is one of the most frequently used forms of figurative language, with most people interacting with hyperbole and extreme exaggerations daily. Whether replying “OMG, I’m literally dying” to a friend’s funny text message or telling a co-worker you “totally recommend” the movie you saw the night before, hyperbole is omnipresent in interpersonal communication. Despite its pervasiveness in written and spoken discourse, hyperbole has received limited empirical attention (Carston & Wearing, 2015; Kreuz et al., 1996; McCarthy & Carter, 2004; Nemesi, 2004; Roberts & Kreuz, 1994; Troiano et al., 2018).

The existing research literature largely studies hyperbole in the context of other forms of figurative speech, such as metaphor and irony, and its communicative function has been assessed primarily through the lens of political rhetoric. As a result, the independent communicative function of hyperbole in casual discourse has yet to receive adequate empirical attention. The current study aims to address this gap in the research literature by using Amazon product reviews to better understand the influence of hyperbole on the perception of review helpfulness. This
research aims to clarify the communicative function of hyperbole, specifically shedding light on its role in persuasive communication strategies in everyday speech.

**Figurative Language**

Figurative language is the use of words or expressions beyond their literal meaning (Glucksberg, 2001; Kalandaze et al., 2018). Tropes, or types of figurative language, including metaphor, irony, and hyperbole, can be valuable tools in everyday discourse, enabling speakers to communicate with greater expressiveness, impact, and memorability (Chakrabarty et al., 2022). However, engaging with figurative language requires listeners to be contextually aware and interpret speakers’ intended meanings, recognizing what it is not meant to be taken literally (Cano Mora, 2009; Carston & Wearing, 2015; Chakrabarty et al., 2022). Likewise, speakers should also consider their audience, being mindful of potential cultural and linguistic nuances that might affect understanding (Alkhammash, 2022; Palmer & Brooks, 2004). Whether it be metaphor, irony, or hyperbole, a shared understanding of both the figurative and literal referent is essential when using figurative language as a communicative tool (Rittio et al., 2021).

Metaphor, irony, and hyperbole are classic examples of figurative language, each fulfilling various discourse goals. Metaphors have received significant attention in the research literature, with some arguing that they are fundamental to our conceptualization of the world (Fussell & Moss, 1998; Lakoff & Johnson, 1980). Linguistically, metaphors play an important role by filling lexical gaps and bridging conceptual divides where literal language falls short (e.g., computer mouse) (Claridge, 2011; Fainsilber & Ortony, 1987). Beyond its lexical role, metaphor is most frequently linked to persuasive communication (Boeynaems et al., 2017; Ottati & Renstrom, 2010; Rossi & Macagno, 2021). When used in persuasive communication, metaphor has been shown to be more effective in producing attitude change than literal language
(Sopory & Dillard, 2002). In addition, speakers gain more terminal credibility, that is, the credibility gained after delivering a message, when using metaphorical persuasion than without (Sopory & Dillard, 2002).

Like metaphor, irony has also received considerable empirical attention. The research literature highlights various communicative functions of irony, including humor (Gibbs et al., 2014; Matthews et al., 2006; Pexman et al., 2005; Roberts & Kreuz, 1994), expression of negative emotion (Roberts & Kreuz, 1994; Wilson & Sperber, 1992), and politeness (Brown & Levinson, 1987; Ivanko et al., 2004). Among these, humor is acknowledged as the primary goal achieved through irony.

Hyperbole stands out as one of the most recognizable and commonly used forms of figurative language (Gibbs, 1994; Stewart & Kreuz, 2003). However, it remains understudied relative to metaphor and irony (Cano Mora, 2009; Claridge, 2011; Kreuz & Roberts, 1993; Zhang & Wan, 2022). The lack of empirical research on hyperbole is likely due to its frequent co-occurrence with other tropes, primarily metaphor (Claridge, 2011; Carston & Wearing, 2015). For example, Claridge’s (2011) study of the Santa Barbra Corpus of Spoken American English (SBC) and the British National Corpus (BNC) revealed that between fourteen and twenty percent of hyperbole includes metaphor (Claridge, 2011; Burgers et al., 2018). Further, Hsiao and Su (2010) found that hyperbole and metaphor often co-occur in the case of affective communication, suggesting metaphor serves as the foundation for hyperbole (Burgers et al., 2018). Likewise, Kreuz et al. (1996) conducted a literary corpus analysis on contemporary literature and found that not only is hyperbole used most in combination with other tropes, but most of hyperbole’s co-occurrences involve metaphor. Many argue that this co-occurrence
demonstrates hyperbole as a subtype of metaphor or irony, resulting in the neglect of hyperbole as a subject of independent study (Carston & Wearing, 2015; Gibbs, 2000).

**Hyperbole**

Hyperbole, as defined by the *Oxford English Dictionary*, uses exaggerated statements to convey intense emotions, with the understanding it is not to be taken literally (Oxford University Press, n.d.). This definition is supported by the research literature, highlighting the deliberate use of overstatement, which occur when a speaker intentionally makes excessively strong claims by using extreme language (e.g., “I stood in line for a million years”), ultimately surpassing what is typically warranted by the subject (Burgers, Brugman, et al., 2016; Cano Mora, 2006; Carston & Wearing, 2015; Norrick, 2004; Zhang & Wan, 2022). As with all forms of figurative language, hyperbole can only be used successfully when the listener recognizes the distinction between a message’s literal use of language and what is being implied figuratively (Claridge, 2011; Henkemans, 2013).

In addition, emotional expression and evaluation are important functions of hyperbole (Carston & Wearing, 2015). Claridge (2011) distinguishes hyperbole as a linguistic intensifier that can be used to evaluate an object or target. For example, if someone is pleased with a pair of shoes, they might say, “these shoes are my *holy grail,*” indicating their extreme liking for the shoes beyond what is warranted in a literal sense. Hyperbolic evaluation operates on a degree scale, with judgments becoming more severe as the difference between the literal and hyperbolic expression increases.

Building on the evaluative function of hyperbole, Claridge (2011) adds that speakers communicate emotional states or attitudes through the hyperbolic evaluations. Emotional evaluation can be accomplished through valence-specific hyperbole, where the valence of a
hyperbolic expression reveals the underlying emotional value (e.g., positivity or negativity) associated with the hyperbolic term (Barrett & Russell, 1999; Lang et al., 2013). The degree of discrepancy between the literal and hyperbolic expressions, combined with the hyperbolic term valence, reflects the strength and direction of a speaker’s emotion (Claridge, 2011). Emotional engagement is central to hyperbole, and through exaggerated language, speakers communicate their “emotional truth” (Claridge, 2011, p. 20).

Although metaphor is widely associated with persuasion, the research literature’s well-documented co-occurrence between metaphor and hyperbole provides anecdotal support for hyperbole’s role in persuasion (Claridge, 2011; Henkemans, 2013; Nemesi, 2004). Persuasion involves influencing one’s mental state before subsequent behavioral changes occur (O’Keefe, 2002), with a widely accepted definition as "the active attempt by an individual, group, or social entity to change a person’s beliefs, attitudes, or behaviors by conveying information, feelings, or reasoning" (Cacioppo et al., 2018, p. 129). According to Cacioppo et al.’s (2018) definition, speakers should focus on how they communicate information or feelings to increase their persuasive power.

Hyperbole’s role as a linguistic maximizer naturally draws attention to itself and evokes emotion, highlighting the significance and relevance of a message (Claridge, 2011). This heightened saliency increases a message’s impact on the listener and strengthens the speaker’s argument, both of which are important for achieving persuasive communication (Claridge, 2011; Ferré, 2014; Henkemans, 2013). In addition, Claridge (2011) connects hyperbole’s emotive quality to Aristotle’s theory of pathos, an ancient rhetorical device introduced in Greek philosophy as a method of persuasion. When using pathos for persuasive communication, a speaker leverages language to evoke specific emotions from the listener in hopes of influencing
their thoughts, opinions, or judgment, making them more receptive to the speaker’s argument (Rapp, 2023). Hyperbole’s use of emotional exaggeration can be applied, as pathos suggests, to alter a message’s perception of importance, urgency, desirability, or intensity (Claridge, 2011; Rittio et al., 2021).

When used in persuasive communication, hyperbole allows speakers to communicate information, feelings, and reasoning, surpassing the ability of literal language, all while maintaining brevity (Henkemans, 2013; Rittio et al., 2021). Additionally, hyperbole’s ability to increase message saliency and draw attention to a speaker’s argument, while also evoking and arousing emotion in the listener through emotionally charged evaluative language, further justifies why persuasion should be considered and investigated as a potential communicative function of hyperbole (Henkemans, 2013).

The co-occurrence of hyperbole and metaphor, the extensive research literature on the persuasiveness of metaphor, and the research literature discussed here suggests a strong relationship between hyperbole and persuasion (Claridge, 2011; Henkemans, 2013; Nemesi, 2004; Rittio et al., 2021). While the interaction between hyperbole and persuasive communication has been explored, the empirical research is limited and primarily applied in political discourse (Alattar, 2017; Burgers, Konijn et al., 2016; Rausser et al., 2020). The study of hyperbole and persuasion in everyday speech is still in its early stages, with only a handful of empirical studies exploring this topic in an casual context (Christodoulidou, 2011; McCarthy & Carter, 2004).

Politicians frequently use persuasive hyperbole to bring attention to a topic, increase public interest, and elicit emotional responses from the public to shape how their message is perceived, framing them in a way that aligns with the politician’s agenda (Burgers, Konijn, et
al., 2016; Craig & Blankenship, 2011). Deflection of criticism is an example of this (Abbas, 2019; Hodges, 2020). Politicians may use hyperbole to deflect criticism or distract from inconvenient truths by using exaggerated language to divert attention away from substantive issues or weaknesses in their arguments (Abbas, 2019; Hodges, 2020). Hyperbolic expressions are used in political campaigning to energize supporters by making bold promises or painting stark contrasts between themselves and their opponents (Weber & Wirth, 2014).

Burgers, Konijn et al. (2016) introduce figurative framing theory, proposing that figurative language tropes serve as both framing devices, providing linguistic cues (tone, cadence, stress) and as reasoning devices (containing important conceptual content). According to the authors’ theory, hyperbolic frames function as reasoning devices, directing and shaping political discourse by moderating the type or valence of a discussion surrounding a given topic.

Alatter (2017) used three of former President Barack Obama’s speeches from the 2012 presidential campaign to extract and analyze 181 hyperbolic expressions to better understand the production process of hyperbolic language in political discourse. The results revealed emphasis, evaluation, and vagueness as the communicative functions of hyperbole in political communication.

Similarly, Rausser et al. (2020) looked at the polarization between parties within political discourse and the role of hyperbole in politicians’ battle for narrative control. The results reinforced Burgers, Konijn et al. ’s (2016) conclusion, reaffirming the advantage that comes with controlling the narrative is a result of unconstrained use of hyperbole, allowing for public opinion to be directed in a preferred direction.
In the context of casual discourse, empirical research focused on hyperbole is limited. However, Christodoulidou’s (2011) and McCarthy and Carter’s (2004) research provide an introductory look into the role of exaggerated language in everyday, informal language.

Christodoulidou (2011) examined the use of hyperbole in casual, natural Cypriot Greek conversations between family and friends. The results emphasized the interactive nature of hyperbole, supporting intensification, humor and banter, solidarity, antipathy, and intimacy as communicative functions of hyperbole in everyday speech (Christodoulidou, 2011). Moreover, Christodoulidou (2011) found that hyperbole is often used in self-deprecation and the exaggeration of real-life situations, characteristics, or events – referred to as upscaling reality – to make them appear more significant, dramatic, or extreme (Christodoulidou, 2011).

In addition, McCarthy and Carter (2004) examined a 5-million-word spoken English corpus, looking at the deliberate use of hyperbole in everyday British English speech. The results of their linguistic analysis revealed persuasion and evaluation as the two main functions of hyperbole in casual conversation (McCarthy & Carter, 2004).

While Christodoulidou (2011) and McCarthy and Carter (2004) offer empirical support for hyperbole’s prevalence in everyday speech, the average person can observe hyperbolic expressions in casual conversations. For example, consider you are trying to convince your friend to watch your favorite movie, but they are hesitant and not sure they will enjoy it. You may try to convince them by describing the movie as "life-changing," "outstanding," or "the best of all time." According to Cacioppo et al. (2018), you are attempting persuasive communication using hyperbolic expressions.

Empirically, the study of hyperbole and persuasion in everyday speech is still in its early stages, with limited supporting research literature (Christodoulidou, 2011; McCarthy & Carter,
However, there are abundant opportunities for future research, such as examining different methods and modalities of everyday speech. For most, using and engaging with hyperbolic language is a daily occurrence (Harman & Strine, 2023). However, there are many other forms of casual communication we frequently participate in but may not always think about, such as online product reviews.

**Online Product Reviews**

Online product reviews, as defined by Mudambi and Schuff (2010), refer to “peer-generated product evaluations posted on company or third-party websites” (Mudambi & Schuff 2010, p. 186). Web-based marketplaces, such as Amazon.com, allow customers to leave open-ended reviews with comments about their experiences or opinions regarding a product, which are then posted on the product’s webpage for other consumers to view and interact with. As online shopping has increased in popularity, consumers have become increasingly inclined to engage in both posting and reading online product reviews, seeking guidance and reassurance before committing to purchase decisions (Chen et al., 2022). Based on a 2023 survey including 8,153 U.S. consumers, 74% of respondents reported using product reviews and ratings to learn about products, especially those they had never purchased (Survey: The Ever-Growing Power of Reviews, 2023). In addition, research has shown that for 93% of consumers, online product reviews affect which products they ultimately purchase, indicating that most online shoppers not only read but rely on online product reviews to make informed purchase decisions (Chen et al., 2022; Vimaladevi & Dhanabhakaym, 2012).

Online marketplaces have recognized the significance of online product reviews for consumers and have implemented consumer feedback-driven helpfulness voting mechanisms. For instance, on Amazon, after each review, users were asked “Was this review helpful to you?”
with a yes or no response option. This helpfulness voting system can be used to calculate and display a helpfulness ratio alongside each review (e.g., 20 out of 25 people found the following review helpful), all of which is represented in Figure 1 (Mudambi & Schuff, 2010).

**Figure 1.**
*Example Online Product Review from Amazon.com*

1 of 1 people found the following review helpful

🌟🌟🌟🌟🌟 Good book, February 17, 2013

By Troy Bryan - See all my reviews

Verified Purchase (What’s this?)

This review is from: Out and About at the Zoo (Kindle Edition)

My son thought the story was good. He liked the funny monkeys. He liked the colorful and fun pictures. Worth the read.

Was this review helpful to you? [Yes] [No]

**Note.** This figure shows an example of an Amazon.com product review with the same format as those in our dataset. The helpfulness proportion is outlined in red at the top of the review, and the helpfulness voting function is outlined in red at the bottom of the review.

Such review helpfulness voting mechanisms rely on consumer feedback to evaluate the utility of a review in guiding their purchasing decisions (Yin et al., 2020). Using consumer review-based feedback, these mechanisms assist customers in efficiently navigating thousands of product reviews, enabling them to quickly locate the most helpful reviews and streamline their decision-making process (Mudambi & Schuff, 2010).
When a consumer indicates that a review was helpful, they are not indicating anything about their purchase decision. Rather, helpfulness votes indicate the diagnosticity of a review, that is, a review’s ability to influence the reader’s attitudes and opinions about a given product, demonstrating its perceived value in the decision-making process (Kempf & Smith, 1998; Mudambi & Schuff, 2010). Consider also Cacioppo et al.’s (2018) definition of persuasion which highlights “[changing] a person’s beliefs, attitudes, or behaviors by conveying information, feelings, or reasoning” (p. 129). We agree with the research literature’s finding of helpfulness votes as an indication of attitude and opinion shift and connect it to the widely accepted definition of persuasion offered by Cacioppo et al. (2018) to ultimately support helpfulness as a measure of persuasion in the present study.

Due to the absence of facial expressions and vocal nuances in computer-mediated communication, such as online product reviews, consumers rely on a wide range of linguistic devices to express and translate their opinions, experiences, and evaluations of products (Aerts et al., 2017; Hong & Park, 2012). Likewise, as a reader, consumers are also reliant on numerous review characteristics to decipher which reviews are helpful.

Linguistic analysis research on the determinants of perceived review helpfulness is paramount in the business and marketing literature, offering insights into factors shaping the perceived utility of reviews (Choi & Leon, 2023; Li et al., 2020; Siering et al., 2018; Zhang, 2008). Within this research literature, Amazon is often used to source lexical data. There are several reasons for this, the first being Amazon’s global reach, which helps to ensure diverse language styles within the review text data. In addition, Amazon offers an extensive, diverse range of products, providing a generous sample size of text data for linguistic analysis. Amazon product reviews are also publicly accessible and free to access, making them ideal candidates for
research, such as ours, which benefits from the ability to collect sample datasets for data preprocessing and variable creation.

**Determinants of Review Helpfulness**

Recent studies have examined diverse factors, including review length (e.g., word count, words per sentence), review content (e.g., readability, emotional expression, extremity), and peripheral elements (e.g., product star ratings, reviewer reputation, reviewer identity). In addition, much of the research literature considers the moderating effect of product type, that is, search versus experience goods, on the perceived helpfulness of a review (Chua & Banerjee, 2016; Lee & Choeh, 2016; Mudambi & Schuff, 2010; Park, 2018). Originally proposed by Nelson (1970; 1974), products can be distinguished as either search goods (i.e., cellphones, televisions, cameras), which consumers can assess for quality before purchase, or experience goods (i.e., grocery products, video games, movies), which must be purchased before quality or utility can be assessed.

As a moderating factor, consumers are likely to identify different review characteristics as helpful for search goods compared to experience goods. For example, search good reviews tend to be objective, using instrumental language to discuss product specifications and functionality (Pan & Zhang, 2011; Strahilevitz & Myers, 1998). However, experience good reviews are inherently subjective, often using emotional language to reflect the review author’s personal experience with the product (Pan & Zhang, 2011). The type of language and review content that is perceived as helpful for one may have the opposite effect for the other, making this product distinction important to consider when analyzing predictors of review helpfulness (Ghose & Yang, 2009; Wang et al., 2023).
Mudambi and Schuff (2010) used Amazon product review datasets to study the effects of review extremity (i.e., star rating) and review length (i.e., word count) on review helpfulness, also considering the moderating effect of product type (i.e., search and experience goods). The results demonstrate differing effects of review extremity and length on helpfulness, depending on the type of product being reviewed (Mudambi & Schuff, 2010). For experience goods, review extremity had a negative effect on helpfulness, but not for search goods. Moreover, review length had a more substantial positive effect on helpfulness for search goods than experience goods.

Cao et al. (2011) used latent semantic analysis to examine the effects of a review’s basic (i.e., direct review observations), stylistic (i.e., word count, sentence length), and semantic (i.e., the substance of a review) features on the number of helpful votes a review receives. In extracting the semantic review characteristics, the authors found the substance of a review (e.g., the relevancy and readability) to have the most influence in terms of the number of helpful votes a review received (Cao et al., 2011). In addition, the findings demonstrate a positive impact of review extremity on review helpfulness. That is, reviews expressing extreme opinions received more helpful votes than reviews expressing diverse or neutral opinions (Cao et al., 2011).

Ahmad and Laroche (2015) analyzed the framing effect of discrete emotions (hope, happiness, anxiety, and disgust) to determine their effects on review helpfulness when expressed in Amazon product reviews. The results revealed a positive effect of happiness, stronger than that of hope, on the helpfulness of a review. In addition, anxiety was shown to have a negative effect on review helpfulness, while disgust had a positive effect.

Liu and Park (2015) analyzed tourism review data, looking at the reviewers’ characteristics (e.g., disclosure of personal identity, reviewer expertise, and reputation),
quantitative (e.g., star rating and review length) and qualitative (e.g., readability) review characteristics to identify elements of online reviews that affect review helpfulness. The results revealed that a combination of reviewer and review characteristics had a positive impact on perceived review helpfulness. Specifically, they found review readability and review length to be significant predictors of review helpfulness.

In line with other research, Chua and Banerjee’s (2016) study looked at the effects of review sentiment (e.g., star ratings) and product type (e.g., search versus experience goods) on review helpfulness. Analyzing Amazon reviews from six products, half search goods and half experience, the results revealed that review helpfulness varies as a function of review sentiment, regardless of product type.

Lee and Choeh (2016) also used Amazon product reviews to better understand and identify the product (i.e., number of reviews for a given product), source (i.e., disclosure of reviewer identity, reviewer reputation, and reviewer rank), and content-based (i.e., review extremity and depth) review features that contribute to perceived review helpfulness. In addition, the authors explored the moderating effect of product type (e.g., search versus experience goods). The results indicated a positive effect on review helpfulness for reviewer identity disclosure and review depth. However, product type was shown to moderate these relationships. These include the number of reviews a product has and reviewer identity disclosure that have a greater impact on helpfulness for experience goods, whereas reviewer reputation, extremity, and depth were more impactful for search goods.

Similarly, Park’s (2018) exploratory analysis looked at several linguistic (e.g., star rating, word count, words per sentence, comparative language) and psychological (e.g., positive emotion, negative emotion, cognitive processing, analytical thinking, and perceptual language)
characteristics of Amazon product reviews. Park (2018) examined five products, following Nelson’s (1970; 1974) search and experience goods framework, to uncover their effects on review helpfulness. The results revealed that star rating, word count, and analytical thinking affected the helpfulness of a review across all five products.

**Hypotheses**

This study aims to explore the role of hyperbole in consumer perception of review helpfulness for online product reviews on Amazon.com. We propose that when used in everyday speech, hyperbole elicits emotion and increases message saliency, strengthening a message’s persuasive power. Therefore, we hypothesize that Amazon product reviews containing hyperbole will be positively correlated with a review’s helpfulness ratio, suggesting a positive correlation between hyperbole and persuasive impact.

In addition, we consider valence (e.g., positive and negative hyperbole) and product type (i.e., search and experience goods) as potential moderating factors. In line with the previously reviewed literature, we hypothesize that the valence of a hyperbolic term or expression (e.g., positive and negative hyperbole) may moderate the relationship between hyperbole and perceived helpfulness.

Based on the previous literature addressing the variation of perceived helpfulness according to product type, that is search versus experience goods, and we further hypothesize that product type may moderate the relationship between overall, positive, and negative hyperbole and perceived review helpfulness.
Methods

Hyperbolic Term List Development

To evaluate the effects of hyperbole on review helpfulness, we first established a comprehensive, standardized list of hyperbole to use in our analyses. To do this, we first created a sample review text dataset using Amazon product reviews. Using a random number generator, we selected seven products from Amazon’s best-sellers list. Web scraping software was used to collect the review text data from all seven products. The seven products included sweatpants, an Air Tag, a windshield scraper, purple hair shampoo, sleep headphones, a raspberry baby teether, and a colon cleanse. From each of the seven products, we examined the first 500 reviews, resulting in 3,000 total sample reviews.

After we collected the sample review data, we began the process of independently screening each set of review data for hyperbole. One after another, for each product, each of the three team members independently screened all 500 reviews and review titles, indicating each time a hyperbolic term or phrase was identified. After reading each set of 500 sample reviews, an in-person meeting was held where we discussed and compared our identification of hyperbole throughout the text. The hyperbolic terms and phrases we agreed upon were added to a running list of hyperbole. This process was repeated until all 3,000 sample reviews had been independently screened and then discussed between the three team members. At the end of this process, we were left with a list of hyperbolic expressions. While there were minor discrepancies in our individual identification of hyperbole, the in-person meetings were designed to resolve these differences as a team. When this process was complete, an additional in-person meeting was held to discuss and review each of the items on our hyperbolic term list, to ensure consensus among the three team members.
The final hyperbolic term list contained 249 overall hyperbolic expressions (refer to the Appendix the complete list of terms and phrases, categorized by valence). It is important to note that nine of the 249 terms on the list of hyperbolic expressions, or 3.6%, could not be included due to a programming error within the linguistic analysis software, LIWC (Linguistic Inquiry and Word Count). This error resulted in the exclusion of expressions that started with a number (e.g., “100% recommend”). As a result, 243 terms and phrases remained in the finalized list of hyperbole.

**Term Valence Categorization**

As a team, we took the hyperbolic term list a step further, and separated the term list into three groups according to their valence: positive, negative, and contextual. Positive (e.g., perfect, flawless, magical) and negative (e.g., horrendous, unbearable, worthless) hyperbolic expressions can easily be sorted as the two are objectively different. However, terms that were neither positive nor negative could assume either valence depending on context. Therefore, these terms were sorted into a contextual category (e.g., crazy, OMG, literally). We chose to label this group as "contextual" rather than "neutral" as our third category because hyperbole inherently lacks neutrality (Claridge, 2011).

The hyperbole valence categorization process followed a similar methodology to that of the overall hyperbolic term list. Each team member received a copy of the overall hyperbolic term list to independently review and categorize by valence. Each team member was instructed to categorize each term as being positive, negative, contextual, or unknown. The optional categorization of “unknown” was to prevent inauthentic or forced categorizations where a term’s valence seemed unclear. After independently categorizing the hyperbolic term list by valence, an in-person meeting was held to compare and discuss our valence categorizations. During this
meeting, any terms labeled as “unknown” were discussed until we agreed on a final
categorization of either positive, negative, or contextual for each “unknown” term. The valence
term lists contained 94 positive, 85 negative, and 61 contextual hyperbolic terms and phrases
before removing the nine terms unable to be processed in LIWC. After removing these terms the
valence lists contained 87 positive, 83 negative, 61 contextual hyperbolic expressions.

**Dataset Description and Preprocessing**

We obtained our dataset from a publicly available Amazon product review corpus, as it has demonstrated its utility in similar studies within the research literature, and is categorically
diverse (He & McAuley, 2016; McAuley et al., 2015). The original dataset contained twenty-
four product categories and 41.13 million reviews collected from Amazon between May 1996
and July 2014. The complete dataset was separated into smaller sub-corpora called 5-core
datasets. These files were smaller in size and contained only the products with at least five
reviews. We obtained and used the 5-core datasets as it established a review baseline within our
datasets while still maintaining a generous sample size.

We selected the following 5-core product category datasets: beauty, cellphone, clothing,
grocery, and video. To ensure a diverse range of product types, we followed Park's (2018) model
for product category selection, which uses Nelson's (1970; 1974) search versus experience goods
framework (Mudambi & Schuff, 2010). Beauty and grocery products are categorized as
experience goods, cellphones are considered search goods, clothing is both a search (e.g.,
designer/name brand clothing) and experience good (e.g., non-branded clothing), and video does
not fall into either category. Its role as a digital product helps to diversify our sample (Park,
2018).
The raw data files were not formatted appropriately for our analyses; therefore we first preprocessed the dataset, creating and coding variables where necessary. The raw data download contained variables such as reviewer ID and name, product ID (e.g., ASIN product code), date and time of review publication, helpful votes, total number of votes, review text, review title, and star rating (one through five). For this study, we were only interested in the number of helpful votes, total number of votes, and the review text and title.

The dataset provided the number of helpful votes and the total number of votes for each review, which we used calculate our “review helpfulness” variable (number of helpful votes / total number of votes). In addition, to test our product type hypothesis, we created the product type (search and experience goods) variable and dummy code it using the following nominal variables: search goods = 1; experience goods = 2; both (search and experience) = 3; neither = 4. The product type dummy codes were then added to our dataset and used as filters, allowing us to run independent regressions for search and experience goods.

The original dataset contained 859,998 reviews with the following number of reviews in each product category: beauty, 198,503; cellphone, 194,440; clothing, 278,678; grocery, 151,255; and video, 37,127. However, in line with the research literature, we deleted all cases where there were fewer than ten helpfulness reviews (Mudambi & Schuff, 2010). This resulted in a final dataset of 22,591 total reviews comprising 5,144,118 words (excluding review titles). See Table 1 for the final number of reviews in each product category, after selection and preprocessing.
In addition, many of the variables in our output had large, positively skewed distributions due to the size of the review dataset and the limited size of our hyperbolic term lists by comparison. To correct for skewness, a base ten log transformation was applied to overall, positive, and negative hyperbole. The transformed variables are labeled “Log” to specify this transformation.

In addition, due to the nature of the sample, a significant portion of the dataset contained zeros (i.e., reviews not containing any of our 243 hyperbolic terms and phrases). Reviews not containing hyperbole (i.e., the terms from our list) were not useful in the present study. While our list of hyperbolic expressions was an exhaustive representation of the hyperbolic language in a sample of 3,000 Amazon product reviews, it is only a fraction of the hyperbole that exists in everyday speech. The review text in this study was screened for hyperbole using our list of hyperbole as the criteria, and because our list does not contain every hyperbolic expression imaginable, it is inevitable that some of the reviews in our dataset would inaccurately receive a

**Table 1**

*Description of Datasets after Selection and Preprocessing*

<table>
<thead>
<tr>
<th>Product Category</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beauty</td>
<td>6,490</td>
</tr>
<tr>
<td>Cellphone</td>
<td>4,243</td>
</tr>
<tr>
<td>Clothing</td>
<td>6,494</td>
</tr>
<tr>
<td>Grocery</td>
<td>4,481</td>
</tr>
<tr>
<td>Video</td>
<td>883</td>
</tr>
</tbody>
</table>

*Note.* This table provides the number of reviews in each of the five product categories after preprocessing.
hyperbole score of zero. There is no way to determine if a hyperbole score of zero is a true zero or not. As a result, we filtered out all reviews with a hyperbole score of zero prior to running each regression.

LIWC

Linguistic Inquiry and Word Count (LIWC) is a text analysis software developed by James Pennebaker and colleagues at the University of Texas. LIWC offers various applications, but it is primarily used to conduct linguistic analyses by categorizing language in a given text to uncover and quantify underlying psychological, social, and behavioral phenomena (Boyd et al., 2022; Chung & Pennebaker, 2007; Pennebaker et al., 2015). LIWC is well-established and has demonstrated its utility within the research literature, with several studies incorporating it in their analyses (e.g., Ahmad & Laroche, 2016; Park, 2018; Wang et al., 2019; Yang et al., 2015).

We used LIWC in two ways in this study, the first being to create hyperbole dictionaries. Using LIWC's custom dictionary workbench, we developed four dictionaries, each representing one of the four hyperbole valence lists: overall, positive, negative, and contextual. The overall, positive, and negative hyperbole dictionaries were used by LIWC to scan the product review datasets, tallying occurrences of the hyperbolic expressions from each of the respective dictionaries. LIWC’s dictionary analysis then divided the number of hyperbolic terms and phrases identified in a review by the review’s word count to produce a hyperbole proportion (i.e., the percentage of review text belonging to one of the four dictionaries) for each review.

The second application of LIWC involved extracting embedded emotions (i.e., positive and negative emotions) from our review text using LIWC’s Affect dictionary. This dictionary contains categories that can be used to extract the tone, emotion, anxiety, anger, sadness, and swear words embedded in a body of text. For this study, we were only concerned with extracting
positive and negative emotion terms, using only the subcategory of Emotion from LIWC’s Affect dictionary. According to *The Development and Psychometric Properties of LIWC-22*, p. 11, LIWC’s Affect dictionary includes 337 positive emotion words (e.g., good, love, happy, hope) and 618 negative emotion words (e.g., bad, hate, hurt, tired; Boyd et al., 2022). We used LIWC to automatically read the review text data, comparing each word in a review to the word in the Emotion dictionaries. As a result, LIWC produced a positive and negative emotion score for each review. These emotion scores are a representation of the percentage of total words in a review that aligns with the Positive or Negative Emotion dictionaries.

**Design**

The dependent variable, review helpfulness, is measured by the proportion of people who found a review helpful. Amazon collects review helpfulness data by asking the question “Was this review helpful?” and providing respondents with a “yes” or “no” response choice. We then used both the “yes” and “no” helpfulness votes to calculate the helpfulness proportion, dividing the number of helpful votes by the total number of votes for each review.

The independent variable is overall hyperbole, with positive hyperbole, and negative hyperbole as predictor variables with a moderating effect. Overall hyperbole, positive hyperbole, and negative hyperbole are all measured as a proportion. Each of the three hyperbole proportions are calculated by dividing the number of hyperbolic terms or phrases in a review by a review’s total word count. We also consider the moderating effect of product type (i.e., search and experience goods). See Table 2 for variable specific measurement details.

Tobit regression was used to analyze our data due to the censored nature of our dependent variable, helpfulness. This is because review helpfulness is a continuous variable that is constrained. In other words, consumers cannot indicate the extent to which they found a
review helpful, and as a result, helpfulness cannot be expressed beyond that of perfect helpfulness ratio (i.e., 10 helpful votes out of 10 overall votes). In addition, there is also the potential for a helpfulness bandwagon effect (i.e., the more helpful votes a review has, the more likely it is to be voted as helpful) putting the sample at risk for selection bias. According to Kennedy (1998) and Mudambi and Schuff (2010), this can result in biased estimation, further supporting Tobit regression as the appropriate method of analysis in this study (Ahmad & Laroche, 2015; Filieri et al., 2018; Mudambi & Schuff, 2010, Zhou et al., 2020).
Table 2

Variable Description and Operationalization

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description and Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review Helpfulness</td>
<td>The number of helpful votes a review received (number of helpful votes / numbers of total votes)</td>
</tr>
<tr>
<td>Overall Hyperbole</td>
<td>The number of overall hyperbolic terms and expressions in a review (number of overall hyperbolic</td>
</tr>
<tr>
<td></td>
<td>terms in a review / review total word count)</td>
</tr>
<tr>
<td>Positive Hyperbole</td>
<td>The number of positive hyperbolic terms and expressions in a review (number of positive hyperbolic</td>
</tr>
<tr>
<td></td>
<td>terms in a review / review total word count)</td>
</tr>
<tr>
<td>Negative Hyperbole</td>
<td>The number of negative hyperbolic terms and expressions in a review (number of negative hyperbolic</td>
</tr>
<tr>
<td></td>
<td>terms in a review / review total word count)</td>
</tr>
<tr>
<td>Product Type</td>
<td>1 = search good; 2 = experience good; 3 = both (search and experience); 4 = neither</td>
</tr>
</tbody>
</table>

Note. This table provides the description and operationalization for the dependent variable (i.e., review helpfulness), independent variable (i.e., overall hyperbole), and the variables hypothesized to produce a moderating effect (i.e., product type and hyperbolic term valence).

Tobit regression was used to analyze our data due to the censored nature of our dependent variable, helpfulness. This is because review helpfulness is a continuous variable that is constrained. In other words, consumers cannot indicate the extent to which they found a review helpful, and as a result, helpfulness cannot be expressed beyond that of perfect helpfulness ratio (i.e., 10 helpful votes out of 10 overall votes). In addition, there is also the
potential for a helpfulness bandwagon effect (i.e., the more helpful votes a review has, the more likely it is to be voted as helpful) putting the sample at risk for selection bias. According to Kennedy (1998) and Mudambi and Schuff (2010), this can result in biased estimation, further supporting Tobit regression as the appropriate method of analysis in this study (Ahmad & Laroche, 2015; Filieri et al., 2018; Mudambi & Schuff, 2010, Zhou et al., 2020).

Results

Using Tobit regression models, we examined the effect of overall, positive, and negative hyperbole on review helpfulness as well as the moderating effect of product type (i.e., search and experience goods) on the relationship between overall, positive, and negative hyperbole on review helpfulness. We ran separate regression analyses for each of the predictor variables (i.e., overall, positive, and negative hyperbole), because the data are nested within one another and are not independent.

The Effect of Hyperbole on Review Helpfulness

Our first hypothesis addressed the effect of overall hyperbole on review helpfulness. A Tobit regression analysis was performed with overall hyperbole as the predictor variable. As shown in Table 3, the results demonstrate a statistically significant positive effect of overall hyperbole on review helpfulness, $p < .001$. This implies that as the amount of overall hyperbole in a review increases review helpfulness also increases.

The Moderating Effect of Term Valence

The second hypothesis is that hyperbolic term valence (i.e., positive versus negative hyperbole) may have a moderating effect on the relationship between hyperbole and review helpfulness. To test this hypothesis, two separate Tobit regression analyses were performed, the first with positive hyperbole as the predictor variable. The results also show a statistically
significant positive effect of positive hyperbole on review helpfulness, \( p < .001 \). As with overall hyperbole, these results suggest that as the amount of positive hyperbole in a review increases, so does review helpfulness.

The second regression was run with negative hyperbole as the predictor variable. As Table 3 shows, there is a significant negative effect of negative hyperbole on review helpfulness, \( p = .005 \). The negative coefficient and significant effect imply that as the amount of negative hyperbole in a review increases, the helpfulness decreases.

**Table 3**

*Tobit Regression Output: Effect of Overall, Positive, and Negative Hyperbole on Perceived Helpfulness*

<table>
<thead>
<tr>
<th>Variable</th>
<th>( N )</th>
<th>Coefficient</th>
<th>SE</th>
<th>z Value</th>
<th>Sig.</th>
<th>Log Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Hyperbole</td>
<td>14,028</td>
<td>.023</td>
<td>.006</td>
<td>3.964</td>
<td>&lt; .001**</td>
<td>3323.489</td>
</tr>
<tr>
<td>Positive Hyperbole</td>
<td>10,136</td>
<td>.041</td>
<td>.006</td>
<td>6.302</td>
<td>&lt; .001**</td>
<td>2749.710</td>
</tr>
<tr>
<td>Negative Hyperbole</td>
<td>2,643</td>
<td>-.048</td>
<td>.017</td>
<td>-2.786</td>
<td>.005**</td>
<td>653.483</td>
</tr>
</tbody>
</table>

Lower bound: None. Upper bound: 1.00

*significant at \( p < 0.05 \), **significant at \( p < 0.01 \)

*Note.* This table shows the effect of overall, positive, and negative hyperbole on review helpfulness. \( N \) represents the number of reviews containing the respective type of hyperbole.
The Moderating Effect of Product Type

The third hypothesis addresses the moderating effect of product type (i.e., search and experience goods) on the relationship between overall, positive, and negative hyperbole and review helpfulness. To test this hypothesis, we ran six separate Tobit regression analyses in two stages. The results of these analyses are shown in Table 4.

In the first stage, we used the dummy variables coded for search and experience goods to filter out all review data product type combinations (i.e., experience goods, both search and experience goods, and neither) other than pure search goods. Cases were filtered by product type. Three individual regressions were performed on the search good product reviews, one for overall, positive, and negative hyperbole. The results do not show a statistically significant effect of overall hyperbole on review helpfulness for search goods, $p = .057$. Likewise, there was not a statistically significant effect for positive hyperbole on review helpfulness for search goods, $p = .356$. However, the results did show a statistically significant negative effect of negative hyperbole on review helpfulness for search goods, $p < .001$. This indicates that as negative hyperbolic review language increases, the perceived helpfulness of a review for search goods decreases.

The same process was conducted for experience goods, filtering out all review data for product type combinations (i.e., search goods, both search and experience goods, and neither) other than pure experience goods. Three additional regressions were performed on the experience good review data, one for overall, positive, and negative hyperbole. The results showed a statistically significant effect of overall hyperbole on review helpfulness for experience goods, $p = .049$. This indicates that as positive hyperbolic review language increases, the
perceived helpfulness of a review for experience goods increases, \( p < .001 \). Finally, the effect of negative hyperbole on review helpfulness for experience goods was not significant.

Table 4

*Tobit Regression Output: Effect of Product Type on the Relationship Between Hyperbole and Perceived Helpfulness*

<table>
<thead>
<tr>
<th>Variable</th>
<th>( N )</th>
<th>Coefficient</th>
<th>( SE )</th>
<th>( z ) Value</th>
<th>Sig.</th>
<th>Log Likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Hyperbole * Search Good</td>
<td>2,904</td>
<td>-.024</td>
<td>.012</td>
<td>-1.907</td>
<td>.057</td>
<td>1003.649</td>
</tr>
<tr>
<td>Overall Hyperbole * Experience Good</td>
<td>6,802</td>
<td>.016</td>
<td>.008</td>
<td>1.972</td>
<td>.049*</td>
<td>1813.377</td>
</tr>
<tr>
<td>Positive Hyperbole * Search Good</td>
<td>2,182</td>
<td>-.012</td>
<td>.013</td>
<td>-.922</td>
<td>.356</td>
<td>887.114</td>
</tr>
<tr>
<td>Positive Hyperbole * Experience Good</td>
<td>4,773</td>
<td>.040</td>
<td>.010</td>
<td>4.049</td>
<td>&lt; .001**</td>
<td>1386.515</td>
</tr>
<tr>
<td>Negative Hyperbole * Search Good</td>
<td>754</td>
<td>-.113</td>
<td>.034</td>
<td>-3.327</td>
<td>&lt; .001**</td>
<td>212.741</td>
</tr>
<tr>
<td>Negative Hyperbole * Experience Good</td>
<td>1,209</td>
<td>-.003</td>
<td>.023</td>
<td>-.141</td>
<td>.888</td>
<td>377.305</td>
</tr>
</tbody>
</table>

Lower bound: None. Upper bound: 1.00

*significant at \( p < 0.05 \), **significant at \( p < 0.01 \)

*Note.* This table shows the moderating effect of product type (i.e., search and experience goods) on the individual relationship between overall, positive, and negative hyperbole and review helpfulness. \( N \) represents the number of reviews included in each respective regression after all zeros had been filtered out from the respective hyperbole term variation (i.e., overall, positive, and negative).
Correlational Analysis: Hyperbolic Term Valence and LIWC Emotion Scores

To validate the positive and negative valence hyperbole term lists, we used LIWC’s Positive and Negative Emotion scores to conduct two Pearson correlation analyses assessing the relationship between our hyperbolic term valence categorization and LIWC’s emotion scores. This decision was motivated by the hyperbole emotion literature, which suggests that emotion is an integral part of hyperbole. Therefore, we predicted that the positive and negative list of hyperbole should be positively correlated with LIWC’s Positive and Negative Emotion dictionaries.

The first correlation analysis concerned the relationship between the list of positive hyperbole and LIWC’s positive emotion scores. The results showed a statistically significant relationship between positive hyperbole and positive emotion, $r(22,590) = .137, p < .001$. In the context of online reviews, these results indicate that as the amount of positive hyperbole in a review increases, so does the amount of positive emotion (see Table 5).

The second correlation analysis examined the relationship between the list of negative hyperbole and LIWC’s negative emotion scores. The results showed a statistically significant relationship between negative hyperbole and negative emotion, $r(22,590) = .171, p < .001$. This suggests that as the amount of negative hyperbole in a review increases, so does negative emotion (see Table 5). The results of the correlation analyses provide external validation for the categorization of positive and negative hyperbole, supporting the emotional influence of hyperbolic expression.
Table 5
Correlation Between Positive and Negative Hyperbole and LIWC Positive and Negative Emotion Scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Hyperbole * Positive Emotion</td>
<td>22,591</td>
<td>.137</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Negative Hyperbole * Negative Emotion</td>
<td>22,591</td>
<td>.171</td>
<td>&lt; .001***</td>
</tr>
</tbody>
</table>

***Correlation is significant at the $p < 0.001$ level

Note. This table shows the correlation between the list of positive and negative hyperbole and the Positive and Negative dictionaries developed by LIWC.

Discussion
The research literature on review helpfulness has investigated a number of factors; however, the effect of hyperbole on review helpfulness was previously unexplored. The current study assessed the effect of hyperbole on review helpfulness and the moderating effects of hyperbolic term valence and product type on their relationship.

Overall, hyperbole was shown to have a statistically significant effect on review helpfulness. This is consistent with the research literature, which has found emotional expression in online product reviews to significantly impact perceived review helpfulness (Richins, 1997; Ullah et al., 2015; Yin et al., 2014).

Hyperbolic term valence was also shown to have a significant moderating effect on the relationship between hyperbole and helpfulness, with positive hyperbole having a significant positive effect on helpfulness, and negative hyperbole having a significant negative effect. This is supported by previous research, which has found that reviews containing positive emotion are
more helpful than those containing negative emotion (Chua & Banerjee, 2016; Deng & Ravichandran, 2018; Pan & Zhang, 2011; Ullah et al., 2015).

Product type was also shown to have a significant effect in some cases but not all. In the case of search good reviews, negative hyperbole produced a significant negative effect on helpfulness (i.e., negative hyperbole decreases review helpfulness). However, this effect is consistent with the effect of negative hyperbole on helpfulness when search goods are not included as a moderating variable. Therefore, this does not show a moderating effect of product type (i.e., search goods) on the relationship between negative hyperbole and review helpfulness.

This result is consistent with earlier research, which shows a significant consumer preference for objective style reviews when seeking information regarding a search good (Pan & Zhang, 2011; Strahilevitz & Myers, 1998). The subjective nature of hyperbole explains why there is a negative effect of hyperbole on helpfulness when used in a domain where objective evaluations are preferred. We did not find a significant moderating effect for overall and positive hyperbole in the context of search goods.

For experience goods, overall and positive hyperbole showed a significant positive effect on review helpfulness. This result is supported by the previous research literature which explains the need for experience goods to be purchased before they can be evaluated (Nelson, 1970; 1974; Pan & Zhang, 2011). The literature highlights a consumer preference for subjectively written product reviews for experiential goods, including a reviewer’s emotions, thoughts, and opinions (Ghose & Yang, 2009; Wang et al., 2023). However, we did not find a significant effect of negative hyperbole on helpfulness in the context of experience goods.

It is interesting to note that this study also shows that over 60 percent of the reviews in the dataset contained at least one of the terms from the list of hyperbolic 243 expressions. This
suggests that a relatively small subset of hyperbolic words can have a significant effect on the perceived helpfulness of online product reviews.

**Implications**

The present study is the first to explore the effects of hyperbole on the perceived helpfulness of online product reviews. As a result, the future researchers should consider the potential interaction effects of hyperbole when examining how variables such as emotion, sentiment, and valence affect review helpfulness.

In addition, this study expands on the current understanding of hyperbole in casual discourse and demonstrates its role in persuasive communication. The present study found that when reading online product reviews, consumers may be affected by hyperbole, finding reviews that use hyperbolic expressions more helpful, and potentially more persuasive, than those that do not. This suggests that hyperbole may be used as a tool to increase persuasive power, drawing attention to a speaker’s message and convincing others of its importance.

Finally, in term of the previous work on helpfulness, this study shows the utility of hyperbole in online product reviews. Advertisers and marketers could make use of this finding to increase the helpfulness and possibly the persuasiveness of their messages.

**Limitations**

As with all research, there are limitations of this study that should be considered. First, the review text dataset we used in this study were sourced exclusively from Amazon. This review dataset may not be an accurate representation of the type of language used in the broader online review sample, or the diverse perspectives that exist on other online review platforms such as Yelp or TripAdvisor.
Second, hyperbole was categorized and labeled using broad valence groups (i.e., positive and negative) and validated using LIWC’s large collection of positive and negative emotion words. The categorization and validation of our hyperbolic term list does not provide any insight into the variation in helpfulness that may exist within both the positive and negative valence categories. Therefore, we do not know which of the 243 terms in the list of hyperbolic terms is the most impactful in driving review helpfulness.

**Future Directions**

According to the findings of this study, we have several suggestions for future research. First, to increase the generalizability of this study’s results, future research should investigate the effect of hyperbole on review helpfulness using review data from other e-commerce sites.

In addition, future research should consider refining our list of hyperbole, categorizing terms by their associated discrete emotion or by level of intensity (i.e., are some hyperboles stronger than others?). Previous research demonstrates a significant effect of discrete emotion expression on helpfulness. Applying these concepts to the hyperbole would help clarify the specific characteristics of hyperbole that are perceived as helpful (Ahmad & Laroche, 2015; Srivastava et al., 2024; Yin et al., 2014).

Moreover, one goal of the present research was to examine the moderating effect of term valence (i.e., positive and negative hyperbole) and product type (i.e., search and experience goods), by running individual Tobit regression analyses. While the present study shows that these factors are correlated, a proper moderation analysis of product type (i.e., search and experience goods) and term valence (i.e., positive and negative hyperbole) could be performed to fully explore the effect of these factors on the relationship between hyperbole and helpfulness.
Finally, this study shows that hyperbole and the perceived helpfulness of Amazon product reviews are correlated. However, to better understand this relationship, future research should consider systemically manipulating the presence of hyperbole in Amazon product reviews. This would determine whether a causal relationship exists between hyperbole and review helpfulness.
References


and conversational analysis study of hyperbole in interaction [Unpublished dissertation, University of Valencia].


Chua, A. Y. K., & Banerjee, S. (2016). Helpfulness of user-generated reviews as a function of
review sentiment, product type and information quality. *Computers in Human Behavior, 54*, 547–554. https://doi.org/10.1016/j.chb.2015.08.057


https://doi.org/10.1207/s15327868ms0204_2


https://doi.org/10.1177/1470593107073842


https://doi.org/10.1515/mc-2014-0003


Fussell & R. J. Kreuz, (Eds.). *Social and cognitive approaches to interpersonal communication*. Psychology Press. https://doi.org/10.4324/9781315805917


https://doi.org/10.1080/0144929X.2016.1173099


https://doi.org/10.1016/j.tourman.2014.09.020


https://doi.org/10.1016/j.chb.2017.03.053

Matthews, J. K., Hancock, J. T., & Dunham, P. J. (2006). The roles of politeness and humor in

https://doi.org/10.1207/s15326950dp4101_2


https://doi.org/10.1145/2766462.2767755


https://doi.org/10.2307/20721420


https://doi.org/10.1016/j.jretai.2011.05.002


https://doi.org/10.1080/026999396380079


Yang, Y., Yan, Y., Qiu, M., & Bao, F. (2015). Semantic analysis and helpfulness prediction of text for online product reviews. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on...*


https://doi.org/10.1109/MIS.2008.95

https://doi.org/10.18653/v1/2022.naacl-main.440
Appendix

Complete List of Hyperbolic Terms and Phrases Separated by Valence

Below is the complete list of the hyperbolic terms established in this study, organized alphabetically, and separated by respective valence category. The size of the valence term lists is as follows: 94 positive, 85 negative, 61 contextual, and 249 total hyperbolic terms and phrases. The nine terms omitted from our analysis due to LIWC software issues are denoted by an asterisk.

Positive

#1 Awesome
5/5* Best by far
10/10* 10 star* Best ever
10 stars* Best of all time
100% happy* Best yet
100% recommend* Best thing
100% safe* Beyond excited
A+ Blissful
A must Boom shakalaka
Absolute favorite Brilliant
Absolute must By far the best
Absolutely love Cannot beat this
Absolutely loved Changed my life
Adores, Amazingly Couldn’t be happier
Definitely recommend

Dirt cheap

Dream come true

Essential

Exceptionally

Fab

Fabuloso

Fantastic

Fantastically

Five stars

Flawless

Forevermore

Game changer

Genius

GOAT

God sent

Gorgeous

Grade A

Ideal

Incredible,

Ingenious

Ingeniously

Insane

Invaluable

Killer

Life changing

Lifesaver

Luxurious

Magic

Military grade

Miracle

Must buy

Must have

Must-have

Never been better

No brainer

Obsession

Outstanding

Perfect

Perfection

Perfectly

Perfecto

Praise god

Priceless

Saved my life
<table>
<thead>
<tr>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saving grace</td>
<td>100% dumb*</td>
</tr>
<tr>
<td>Seriously awesome</td>
<td>100% waste of money*</td>
</tr>
<tr>
<td>Soul cleansing</td>
<td>Absolute garbage</td>
</tr>
<tr>
<td>Stunning</td>
<td>Absolute joke</td>
</tr>
<tr>
<td>Super</td>
<td>Absolute trash</td>
</tr>
<tr>
<td>Superb</td>
<td>Absolutely atrocious</td>
</tr>
<tr>
<td>The best</td>
<td>Absolutely hated</td>
</tr>
<tr>
<td>The bomb</td>
<td>Absolutely unacceptable</td>
</tr>
<tr>
<td>The tits</td>
<td>Abysmal</td>
</tr>
<tr>
<td>The truth</td>
<td>Agony</td>
</tr>
<tr>
<td>Thrilled</td>
<td>All hell had broken loose</td>
</tr>
<tr>
<td>Tirelessly</td>
<td>As hell</td>
</tr>
<tr>
<td></td>
<td>Beyond stupid</td>
</tr>
<tr>
<td></td>
<td>Blew it up</td>
</tr>
<tr>
<td></td>
<td>Complete flop</td>
</tr>
<tr>
<td></td>
<td>Complete ripoff</td>
</tr>
<tr>
<td></td>
<td>Complete rubbish</td>
</tr>
<tr>
<td></td>
<td>Complete trash</td>
</tr>
<tr>
<td>Word</td>
<td>Synonym</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Crap</td>
<td>Horrible</td>
</tr>
<tr>
<td>Crappy</td>
<td>Horribly</td>
</tr>
<tr>
<td>Crippled</td>
<td>Horrific</td>
</tr>
<tr>
<td>Debilitating</td>
<td>Horrified</td>
</tr>
<tr>
<td>Destroyed</td>
<td>Junk</td>
</tr>
<tr>
<td>Detests</td>
<td>Miserable</td>
</tr>
<tr>
<td>Devastated</td>
<td>Mockery</td>
</tr>
<tr>
<td>Disastrous</td>
<td>Most egregious</td>
</tr>
<tr>
<td>Egregious</td>
<td>Nails on a chalkboard</td>
</tr>
<tr>
<td>Evil</td>
<td>Nauseated</td>
</tr>
<tr>
<td>Excruciating</td>
<td>Never again</td>
</tr>
<tr>
<td>Fatal</td>
<td>Never ever</td>
</tr>
<tr>
<td>Fatal flaw</td>
<td>Nightmare</td>
</tr>
<tr>
<td>Frantically</td>
<td>No stars</td>
</tr>
<tr>
<td>Freakishly</td>
<td>Nutso</td>
</tr>
<tr>
<td>Garbage</td>
<td>Pain in the ass</td>
</tr>
<tr>
<td>Gory</td>
<td>Piece of junk</td>
</tr>
<tr>
<td>Gut wrenching</td>
<td>Plagued</td>
</tr>
<tr>
<td>Harrowing</td>
<td>Pointless</td>
</tr>
<tr>
<td>Harsh</td>
<td>Poisonous</td>
</tr>
<tr>
<td>Hell</td>
<td>Rubbish</td>
</tr>
<tr>
<td>Horrendous</td>
<td>Shitty</td>
</tr>
<tr>
<td>Horrid</td>
<td>Sucks</td>
</tr>
</tbody>
</table>
Sweating bullet
Terrified
Terrifying
Terrible
Terribly
The worst
Torture
Total gimmick
Total waste
Tragic
Trash
Unbearable

Unusable
Useless
Waste of money
Waste of time
Way too harsh
Worst ever
Worst I’ve ever had
Worst of my life
Worthless
Zero
Zero stars

Contextual
A ton
Absolute
Absolutely
Af
All the time
Bloody
Completely
Constantly
Continuously

Couldn’t be more
Countless
Crazy
Dying
Everything
Extreme
Extremely
Forever
Freaking
Geyser like
Gigantic
Glued to
Hands down
Hard core
Hella
Hilariously
Holy
Huge
Imaginable
Immensely
Impossible
Incredibly
Inevitably
In no time
Insanely
Instantly
Intense
Literally
Massive
Million
Non stop
Oh lord

OMG
One million
Ridiculously
Severe
Severely
Shocked
Shocker
Shockingly
Sickening
Thousand times
Tons of
Tremendously
Truly
Ultra
Unworldly
Wholeheartedly
Wicked
Without a doubt
Wow