Decision-Making Processes and Influential Cues in Social Media

Rezaee Behzad

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT MARKETING

Decision-Making Processes and Influential Cues in Social Media

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Doctorat en Management

par

Behzad REZAEE

Directeur de thèse
Prof. Markus Christen

Jury

Prof. Christian Peukert, Président
Prof. Tobias Schlager, expert interne
Prof. Bruno Kocher, expert externe

LAUSANNE
2023

UNIL | Université de Lausanne

FACULTÉ DES HAUTES ÉTUDES COMMERCIALES

DÉPARTEMENT MARKETING

Decision-Making Processes and Influential Cues in Social Media

THÈSE DE DOCTORAT

présentée à la

Faculté des Hautes Études Commerciales
de l'Université de Lausanne

pour l'obtention du grade de
Doctorat en Management

par

Behzad REZAEE

Directeur de thèse
Prof. Markus Christen

Jury

Prof. Christian Peukert, Président
Prof. Tobias Schlager, expert interne
Prof. Bruno Kocher, expert externe

LAUSANNE
2023

UNIL | Université de Lausanne
Faculté des hautes études commerciales
Internef - 261
CH-1015 Lausanne

IMPRIMATUR

La Faculté des hautes études commerciales de l'Université de Lausanne autorise l'impression de la thèse de doctorat rédigée par

Behzad REZAEE

intitulée

Decision-Making Processes and Influential Cues in Social Media

sans se prononcer sur les opinions exprimées dans cette thèse.

Lausanne, le 02.11.2023

Professeure Marianne Schmid Mast, Doyenne

Thesis Committee

« Prof. Markus Christen »

Full Professor, Université de Lausanne

Thesis Supervisor

« Prof. Tobias Schlager »

Full Professor, Université de Lausanne

Internal Expert

« Prof. Bruno Kocher »

Full Professor, Université de Neuchâtel

External Expert

University of Lausanne
Faculty of Business and Economics

Ph.D. in Management

I hereby certify that I have examined the doctoral thesis of

Behzad REZAEE

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: 31 OCTOBER 2023

Prof. Markus CHRISTEN
Thesis supervisor

University of Lausanne
Faculty of Business and Economics

Ph.D. in Management

I hereby certify that I have examined the doctoral thesis of

Behzad REZAEE

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: October 31, 2023

Prof. Tobias SCHLAGER
Internal expert

University of Lausanne
Faculty of Business and Economics

Ph.D. in Management

I hereby certify that I have examined the doctoral thesis of

Behzad REZAEE

and have found it to meet the requirements for a doctoral thesis.
All revisions that I or committee members
made during the doctoral colloquium
have been addressed to my entire satisfaction.

Signature: _____ Date: 31.10.2023

Prof. Bruno KOCHER
External expert

Acknowledgments

I would like to express my heartfelt gratitude to my dedicated professors, Prof. Markus Christen and Prof. Tobias Schlager. Their unwavering support and invaluable guidance throughout my Ph.D. journey at the University of Lausanne have been instrumental in shaping this thesis. Their constant help and motivation have been the cornerstone of my academic growth. I extend my sincere appreciation to Bruno Kocher, the external member of my thesis committee, for his insightful feedback and contributions that enriched the quality of this work.

I owe a debt of gratitude to my family members: my mother, father, and brother, whose unwavering love, encouragement, and sacrifices have made me who I am today. Their belief in my potential has been my driving force. Lastly, I want to thank all my friends for their support, encouragement, and understanding throughout this academic journey. Your friendship and camaraderie have provided a source of strength and inspiration.

This thesis would not have been possible without the collective support and encouragement of these remarkable individuals, and for that, I am profoundly thankful.

Table of Contents

Synthesis Report – Decision Making Processes and Influential Cues in Social Media

In the current era of digitization, the extensive utilization of social media has revolutionized our communication methods. People can easily share personal stories and opinions, connect with loved ones, obtain news from diverse sources, and follow their preferred brands. In 2023, a remarkable 5.2 billion individuals were reportedly active users of different social media and messaging apps, spending an average of 2.5 hours daily on these platforms[1]. This development has not only enhanced online interaction but also broken down geographical and time barriers, challenging the dominance of traditional media with a new two-way information exchange. This change has opened up possibilities for both individuals and enterprises, redefining how people, as well as brands and consumers, interact. The growing impact of social media on consumers has spurred a significant academic investigation into its effect on our everyday lives. This thesis, comprising three articles, aims to investigate the signals and cues utilized by individuals as they search for information to make informed decisions in the online realm.

The advent of social media has subjected us to a constant influx of information, which, while diverse in purpose, consumes a considerable amount of our time daily. However, this plethora of data does not assure optimal decision-making. As Herbert A. Simon highlighted an excess of information can overwhelm our focus and attention (1971). The human cognitive system, known for its limited processing capacities, is likely to experience compromised decision-making quality under information excess (Gross, 1964). The rise of social media has increased the likelihood of consumers experiencing information overload (Rodriguez et al.,

[1] https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

2014), leading to a compromise between efficiency and accuracy in assessing information credibility (Fogg, 2003). To navigate this, consumers often employ specific cues and heuristics to aid their decisions (Tversky & Kahneman, 1974) and evaluate social media content (Hilligoss & Rieh, 2008; Ranganathan, 2012).

The Elaboration Likelihood Model (ELM) outlines two information processing routes: central and peripheral (Petty & Cacioppo, 1986). The central route involves careful evaluation of information content, which can be inefficient in an information-rich environment. Conversely, the peripheral route employs heuristics, requiring less cognitive effort and aiding decision-making amid uncertainty (Tversky & Kahneman, 1974). This dynamic highlights the challenge of balancing efficient information gathering with effective decision-making on platforms like Twitter, where users typically spend only three seconds per tweet (Counts & Fisher, 2011). In this thesis, we concentrate on the peripheral routes of decision-making on social media, as these appear to be the common shortcuts that users typically employ for their decision-making.

The first article systematically reviews papers centered on decision-making in digital environments. It summarizes the prevailing literature from business, management, and communication journals over the past decade, with a focus on the cues and signals individuals use for decision-making. This includes assessing the credibility and authenticity of users, their social media posts, and the sponsored content influencers produce across various platforms such as Twitter, Instagram, and YouTube. Ultimately, the paper identifies research gaps in the literature and offers practical suggestions for brands and influencers to align content creation more closely with the decision-making strategies of users and consumers.

The second article addresses a critical research gap identified in the first essay, focusing on how people assess the credibility of social media posts, such as tweets, without access to the source of information and relying solely on available post stats (i.e., the number of comments and likes). It investigates a phenomenon known as 'ratioing,' where a post receives more comments than likes. We hypothesized that users find high-ratio posts less credible than low-ratio ones, and we explored and tested this through one field study and five experiments. Findings indicate that high ratios are perceived as a lack of consensus, affecting users' buying behavior and purchase intention. This article enhances our understanding of information processing on social media, offering valuable insights for companies. Specifically, we propose a content creation strategy to mitigate the adverse effects of ratioing, particularly for potentially controversial posts.

The third article examines another research gap identified in the first essay, focusing on how consumers respond to the comment section of posts, especially when it is restricted by the user who made the post. We posit that when social media posts restrict others, such as followers, from commenting, it might negatively affect how consumers view the brand behind the post. Specifically, consumers might see the inability to comment as stifling free speech, leading to unfavorable opinions towards the brand or individual responsible for that post. We used a field study and seven online experiments to validate this hypothesis. The studies reveal that the act of closing the comment section influences attitudes toward the brand through perceived censorship, increasing user suspicion. We also found that this pattern holds true across different scenarios, including the significance of the topic and whether blocking comments is a consistent communication strategy by the author.

Review Paper: Unlocking the Power of Cues and Signals for Informed Decision-Making on Social Media

ABSTRACT

In today's digital age, the widespread use of social media has transformed how we communicate. Individuals effortlessly share personal narratives and opinions, engage with family and friends, access news from various outlets, and keep tabs on their favorite brands. This has not only enriched the online interactive experience but also shattered geographical and temporal boundaries while challenging the information monopoly once held by traditional media, by the new bi-directional path of information sharing. This shift has created opportunities for both individuals and businesses, reshaping the way people connect, particularly for both brands and consumers. The increasing influence of social media on consumers has led to extensive academic research exploring its impact on our daily lives. This article focuses on the field of business and management studies, specifically examining the signals and cues that individuals employ when seeking information for their decision-making, whether in the online or offline realm. These signals facilitate or hinder actions like adopting new ideas, sharing content, and making online purchases. Our goal is to uncover the underlying motivations behind these actions and, in turn, help businesses refine their strategies for effective social media engagement. Additionally, we aim to highlight potential research gaps in the literature that need addressing to better understand the decision-making process of users in social media.

Keywords: Social Media, Information Processing, Decision Making, Social Cues, Heuristics

1. Introduction

Social media have become an integral part of nearly everyone's daily life. Whether it is for entertainment, connecting with friends, staying updated on the latest news, or even as a means of livelihood, social media have permeated our existence. In 2023, an astonishing 5.2 billion people were estimated to be active users of various social media platforms and messaging apps, with an average daily usage of 2.5 hours per person on those platforms[2]. Furthermore, in contrast to traditional media, where individuals were often categorized as either content consumers or content creators, there is now a blurred line between these roles on social media platforms. This is because almost all of us, at some point, have made contributions to social media by posting stories on Instagram, tweeting on Twitter, or engaging with posts through comments and likes. This surge in usage has not only intensified the quantity of information but has also elevated the quality of communication, owing to the two-way interaction between users and content creators, including brands, influencers, politicians, celebrities, and more (Hudson et al., 2016). For example, currently, more than 25 million businesses are leveraging Instagram's robust ecosystem, which includes built-in shopping features, sophisticated advertising platforms that offer user targeting and retargeting options, and various communication tools. These capabilities enable businesses to engage more effectively with both prospective and existing customers, something that was not possible before the rise of social media[3].

On the other hand, we are constantly bombarded with information from various sources, ranging from textual to visual content. This constant stream serves various purposes, whether it is for entertainment, business, or news, and can occupy us for hours each day. However, this onslaught of information cannot guarantee the perfect decision-making for us.

[2] https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/
[3] https://blog.gitnux.com/instagram-shopping-statistics/

Herbert A. Simon's words, "wealth of information creates a poverty of attention" (1971, p. 40), aptly capture this predicament. Essentially, when we are inundated with an overload of information, our ability to concentrate on any single piece of it diminishes. Compounding this challenge is the fact that individuals have a limited capacity to process a vast amount of information, leading to what we commonly refer to as "information overload" (Frias et al., 2008). The rise of social media has only exacerbated this issue, as pointed out by Rodriguez and colleagues (2014) by making the information sources more diverse than ever. Consequently, people now face the daunting task of efficiently managing their attention amidst this information overload (Simon, 1971). Furthermore, they must strike a delicate balance between quickly assessing information for its reliability and taking the time for a more thorough evaluation (Fogg, 2003). For example, when users typically spend only three seconds reading a tweet, the question arises: how can one strike a balance between efficient information gathering and effective decision-making on social media platforms such as Twitter (Counts & Fisher, 2011)?

Before jumping into questions related to decision-making based on information, it is essential to establish a clear understanding of what information itself constitutes. According to the Merriam-Webster Dictionary, information is defined as "knowledge obtained from investigation, study, or instruction." However, to grasp the philosophical essence of this abstract concept, we must explore its various dimensions. As Madden (2000) aptly pointed out, information exhibits four distinct characteristics.

Firstly, in alignment with Merriam's definition, information serves as a representation of knowledge. It possesses the capacity to impart fresh insights into our decision-making process, allowing our cognitive processes to transition from one state to the next. Secondly, information can be gleaned from a variety of environmental stimuli and phenomena, not all of which are intentionally designed to convey a message. Nevertheless, when interpreted

correctly, these stimuli can be informative and reveal concealed cues and signals. For instance, consider the implications of a photograph posted on a social media platform, its number of likes, and the comments it garners. Thirdly, it is crucial to recognize that information is a component of the communication process itself, rather than constituting the entire communication. This dimension is pivotal in understanding the multifaceted nature of information and how its interpretation can vary within different contexts. Lastly, information can be viewed as a commodity transferred from a sender to a receiver, with its value and significance subject to change over time. This dynamic flow of information underscores the fluid nature of its meaning. Now armed with this comprehensive definition, we are prepared to explore how information is utilized in the decision-making process.

The heuristic-systematic model (HSM) and elaboration likelihood model (ELM) provide a comprehensive framework for understanding how individuals process information and change their attitudes based on the central (systematic) or peripheral (heuristic) routes of information processing (Chen & Chaiken, 1999; Petty & Cacioppo, 1986). The central route involves a meticulous examination of the information's content to determine its inherent value before forming an attitude. In essence, information quality becomes the primary determinant of an individual's attitude in this route, which seems not efficient in today's digital era with an abundance of information. On the other hand, the peripheral route relies on peripheral cues or heuristics to shape attitudes, requiring less cognitive effort than the central route. In other words, facing an information overload, consumers often resort to specific cognitive shortcuts known as heuristics, which serve as mental strategies to facilitate decision-making under conditions of uncertainty (Tversky & Kahneman, 1974).

The primary distinction between the two theories lies in how they explain how individuals process information. The HSM suggests that people simultaneously employ two modes, heuristic and systematic, while processing perceived information. On the other hand,

the ELM contends that individuals can utilize only one mode, either peripheral or central, for information processing (Xiao et al., 2018). For example, based on ELM, the extent to which individuals employ information quality, peripheral cues, and heuristics in processing information hinges on their elaboration likelihood, a concept that encompasses both their motivation and ability to evaluate information (Pee, 2012). In essence, it is individuals' willingness and capacity to critically assess the information they encounter on social media that dictate whether they delve deeply into content or rely on shortcuts and cues to navigate the information overload. Another significant discovery from the ELM is that the central route or systematic thinking can result in enduring attitude changes, whereas the peripheral route or heuristic thinking can lead to more transient ones (Petty, 2018).

ELM also sheds light on how decision-making processes differ in digital environments compared to traditional media, such as newspapers and television. As previously mentioned, users on social media platforms may spend only three seconds reading a post, while they are exposed to hundreds of other posts in a single day (Counts & Fisher, 2011). This stands in stark contrast to the days of limited newspaper articles and relatively fewer television channels, where program repetitions were common. Consequently, individuals had more time to ponder newspaper articles or scrutinize the details of TV commercials. According to ELM, users of traditional media had the luxury of dedicating more cognitive effort to these contents. Therefore, they were more inclined to rely on a systematic route of thinking rather than peripheral cues. For instance, factors like source credibility (Hovland & Weiss, 1951) and news medium credibility (Graziano & McGrath, 1986) have always been pivotal in traditional journalism. However, they also relied on characteristic cues, such as the source's identity, familiarity, or design, to assess the information (Metzger et al., 2003; Lucassen & Schraagen, 2013).

In our current digital era, as the information landscape continues to evolve, understanding how individuals navigate the constant stream of information on social media using heuristics provides valuable insights into the dynamics of online decision-making. Consequently, in this article, we aim to illustrate the primary heuristics and cues that consumers routinely employ in their decision-making processes. By doing so, we shed light on how brands can utilize these elements to establish more efficient and seamless communication with both their existing and potential customers.

It's important to note that decision-making on social media extends beyond news consumption and exposure to social media advertisements. It can be relevant even when we are using it purely for entertainment purposes. Take, for example, watching a humorous meme on Twitter. You might decide to interact with the post, leave a comment, share it with your followers, or even follow the user who posted it. These actions tie your digital persona and reputation to the content of the post, particularly the original poster (Metaxas et al., 2015). Thus, decision-making is still present in our interactions, albeit on a potentially smaller scale.

2. Decision-making – central and peripheral

In our daily lives, we are continually exposed to a wealth of information that plays a crucial role in helping us understand our surroundings and make decisions. Even as you read this sentence, your brain is not only processing the text but also taking in environmental cues, perhaps even listening to music, all while making decisions about whether to continue reading. According to the Elaboration Likelihood Model (ELM) proposed by Petty and Cacioppo in 1986, you are essentially following one of two paths in making these decisions.

First, in the context of information processing, one approach is to employ logical and systematic reasoning, a process associated with the central route. This route involves a

structured, objective assessment of the information, often relying on logical causal relationships between the information and the ultimate decision. People typically opt for the central route when they are motivated and capable of engaging in this more effortful cognitive process (Petty, 2013). For instance, consider a question like “What is the capital of Iran?” In this case, you have three options: you may already know the answer, you can take the time to search for it on the internet, or you might choose not to answer the question. In all these cases, you are employing the central route in your decision-making, which, in this context, is aimed at whether or not to answer the question.

Conversely, the peripheral route becomes relevant when there is a lack of motivation or an inability to deeply engage with the information. In such situations, individuals rely on correlational cues and signals to guide their decisions, effectively taking a shortcut that requires less cognitive effort. This route involves drawing associations and patterns from the information rather than delving into the specifics of causality (Petty, 2013). For example, if someone were to ask you, “Which of two Iranian cities is more populated, Tehran or Zanjan?” there is a good chance you might answer “Tehran,” even if you do not have the exact population figures. This tendency to choose Tehran is an application of the recognition heuristic (Goldstein & Gigerenzer, 1999). Essentially, you make a decision shortcut based on your familiarity with the name of Tehran compared to Zanjan, assuming that the better-known city is likely more developed and populated. In this way, you make a decision without possessing the actual answer but by using cues and signals embedded in the question itself, which actually ends up with the correct answer. In the realm of decision-making via the peripheral route, we often rely on various cues, each carrying its own weight in our decision-making process. These cues are typically gathered through specific search algorithms tailored to the situation at hand (Rieskamp & Hoffrage, 1999). What is particularly fascinating is that these cues can exhibit different dynamics with each other depending on the context.

For instance, in some situations, we may employ the cues in a "compensatory" manner. This means that the cues and signals are collected from the context and then weighted based on their perceived importance in our minds, allowing for a comparative evaluation. On the other hand, in different scenarios, the cues function in a "non-compensatory" fashion. This implies that the mere presence of a specific cue can conclude the decision-making process, rendering further search unnecessary (Meinert & Krämer 2022). Furthermore, a heightened level of systematic processing, or availability of systematic information, can reduce the impact of heuristic processing on an individual's evaluation of a message. The concept of the additivity effect highlights that both systematic and heuristic processing independently influence a person's assessment of a message when both modes are present simultaneously (Maheswaran & Chaiken, 1991).

For example, consider a scenario where you are pondering whether Tehran or Zanjan is the capital of Iran, and which is more populous. If your trusted friend, who happens to be an Iranian political scientist, confidently informs you that Tehran holds both distinctions, you may well conclude your search for additional cues in your decision-making process. In this case, the authority and expertise of your friend's opinion act as a powerful non-compensatory cue, outweighing any prior lack of familiarity with Tehran or Zanjan. Thus, as you process the information in this very moment, you may find yourself consciously or unconsciously choosing between these two routes—either engaging in thoughtful, systematic reasoning or relying on shortcuts and cues to make decisions based on the information at hand. Understanding these processes can shed light on how individuals make choices and interpret the vast array of information they encounter in their daily lives.

3. Peripheral decision-making on social media

In today's digital era, the landscape of social media has become an increasingly complex and fast-paced ecosystem, replete with diverse information sources, a myriad of users, brands, and an ever-changing array of content. This complexity often leaves us with little time or inclination to carefully evaluate each piece of information through meticulous, systematic reasoning, commonly referred to as the "central path" in decision-making research (Rodriguez et al., 2014). As a result, we increasingly resort to heuristics and quick cues to navigate this information overload.

For example, in particular, the way we process news information has been fundamentally altered by platforms like Twitter. Unlike the traditional newspaper reading experience, which often employs a more systematic approach to spending time reading and speculating the news, social media encourage users to rapidly assess the credibility of information using heuristic methods (Meinert & Krämer, 2022). These shortcuts might include looking at the authority or popularity of the source, or how widely an article has been shared.

The academic literature on decision-making in social media environments is vast. Many studies focus on either the features of the content itself or the characteristics of the message provider—be it brands, influencers, friends, or regular users. These studies examine how these factors influence consumer behavior, particularly in terms of perceived credibility, authenticity, usefulness, and subsequent purchase intentions. accordingly, research suggests that both the nature of the messenger and the quality of the message play crucial roles in shaping the perceived usefulness of reviews (Liu & Park, 2015).

Source credibility has been also a focal point in understanding persuasion and information processing. While some researchers define source credibility as a feature that influences the audience's perception of the persuasiveness of the speaker (Metzger et al., 2003), others like O'Keefe (1990) argue that it is a receiver-based construct, meaning that the audience grants the speaker a certain level of perceived credibility. Following this line of thought, Wathen and Burkell (2002) concluded that source credibility directly impacts the credibility of the information itself.

The challenge of verifying information credibility is exponentially compounded in the social media context. The absence of strict regulatory oversight, coupled with the bi-directional flow of information, makes it an intricate task requiring a multi-faceted evaluation (Flanagin & Metzger, 2000). Although some existing literature focuses on central routes of perceived credibility, this article aims to explore perspectives related to the elements of Dual Coding Theory.

Dual Coding Theory, first proposed by Paivio, suggests that human cognition operates through two interconnected systems: a verbal system for processing language and a non-verbal system for dealing with visual images. Each system can activate or stimulate the other, allowing for a richer understanding of information when both verbal and visual cues are provided (Paivio, 2013). Hence, in the following sections, we will go deeper into the types of cues—both verbal and non-verbal—that users commonly employ to assess the credibility of both message senders and the information they share on social media and review platforms.

We have organized these cues into four distinct segments (Table 1). The first segment focuses on cues utilized for assessing the quality of online content, such as reviews, comments, and social media posts. The second segment is about the dynamics between online reviews and comments, examining how individuals make decisions when faced with inconsistent information on social media platforms. The third segment investigates the use of

social cues to evaluate the credibility and expertise of content creators, including influencers and bloggers. The fourth and final segment addresses cues employed to assess the validity and reliability of sponsored content on platforms such as YouTube.

It is worth noting that a significant portion of the literature does not solely focus on traditional social media platforms like Twitter and Facebook. It also examines platforms that may not be inherently social media but possess specific features enabling two-way information flow. Examples of such platforms include TripAdvisor and the review sections of websites like Amazon. According to Carr and Hayes (2015, p. 49), social media are defined as "Internet-based, disentrained, and persistent channels of mass-personal communication facilitating perceptions of interactions among users, deriving value primarily from user-generated content." This definition encompasses review platforms as a subset of social media.

However, there is a potential limitation in this definition that may hinder its applicability to research in the context of platforms like Twitter. In review platforms, the flow of information and the network of influence primarily exist between brands and users. In contrast, on traditional social media platforms, the relationships extend beyond brands and users to encompass interactions between users themselves. In our literature review, we have acknowledged this limitation in our research context (Table 1). Moreover, In the upcoming research gap chapter, we will focus deeper into how this limitation may restrict the applicability of research findings to platforms like Twitter and Facebook, for example, when we are dealing with the visibility of the comments in comparison to the reviews.

Table 1
Review of literature on cues and decision-making in the digital environment.

Topic	*Context*	*Major Findings*	*Relevant Literature*
Attitude Toward Online Contents	Platform Reviews (Amazon, TripAdvisor, Booking...)	Reviews play a pivotal role in decision-making, especially on service-based websites, when it comes to purchases.	Santos, 2014; Xie et al., 2014; Tsiakali, 2018; Goh et al., 2013; Liu & Park, 2015
		Votes, rankings, and star ratings on reviews serve as cues for perceived usefulness and also function as filtering tools.	Mudambi & Schuff, 2010; Ghose & Ipeirotis, 2008; Liu & Park, 2015
		The number of reviews can indicate a brand's popularity.	Xie et al., 2014; Chen et al., 2004
	Social Media, and Platform Reviews	Elaboration matters: Longer content is perceived as more useful.	Liu & Park, 2015; Castillo et al., 2011
		Context matters: Longer headlines tend to be viewed as less credible.	Zhang et al., 2015
		Grammar and Punctuation matter: it can increase the perceived credibility of online content.	Morris et al., 2012
		Framing matters: Detailed, specific messages (low construal) are seen as more credible than abstract, generalized messages (high construal).	Reczek et al., 2018
		Multimedia elements matter: Photos from other customers are generally seen as more credible than those from the company itself. However, it can only increase the perceived credibility only if users engage with them.	Filieri, 2015; Li & Sundar, 2022
Dynamics of Online Reviews and Comments	Social Media, and Platform Reviews	Negative comments in isolation can damage a brand's reputation and equity. Moreover, negative comments/reviews are more impactful than positive ones.	Beneke et al., 2016; Royo-Vela & Casamassima, 2011
		Constructive responses to negative reviews can counteract the negative impact and enhance trustworthiness.	Könsgen et al., 2018

Topic	*Context*	*Major Findings*	*Relevant Literature*
		Tie strength matters while reading negative comments/reviews: In-group (strong tie): acts as a diagnostic function, directing to the issue. Out-group (weak tie): can increase the purchase intention.	Bitter & Grabner-Kräuter, 2016; Wen et al., 2009; De Maeyer, 2012
Attitude Toward Online Content Generators	Social Media, and Platform Reviews	Real profile pictures are seen as more credible than Avatars.	Riedl et al., 2014
	Social Media Platforms	Identity cues, authenticity cues, and reputation cues can increase the perceived credibility and perceived usefulness of the content.	Mackie et al., 1990; Sundar, 2008; Özbölük & Akdoğan, 2022
		Reputation cue is more effective when the reputation of the content provider (i.e., influencer) is aligned with the content.	Metzger et al., 2010; Martínez-López et al., 2020
		The verification badge can increase the perceived trustworthiness of the users, but not the credibility of their content.	Vaidya et al., 2019
		The follower-to-following ratio can increase the perceived credibility of the users.	Valsesia et al., 2020
Attitude Toward the Sponsored Content	Social Media Platforms	Disclosure matters: If it happens automatically, it may increase the perceived trust and attitude toward the content. If it happens by the users, the effect is reversed.	Sah et al., 2018
		Intention matters: Brand attitude and purchase intentions can dip when consumers are directed to a product sales page instead of a generic starting page.	Stubb & Colliander, 2019; Vrontis et al., 2021
		The tone of the voice matters: Extremely positive or negative reviews often lead to suspicions of their authenticity.	Moon et al., 2019; Ott et al., 2013

4. Attitude toward online content – verbal and non-verbal

The interactive nature of social media has amplified the significance of user-generated content, making it a key area for analysis in the literature on marketing management. This is particularly true when it comes to online reviews and ratings. For example, according to Santos (2014), almost half of online consumers actively read and post reviews after trying out service products. Moreover, studies reveal that 53% of travelers consult online reviews before finalizing a hotel reservation, and 77% often rely on these reviews when choosing a hotel (Xie et al., 2014). Online reviews serve as a form of electronic word-of-mouth (eWOM), defined as any online statement made by customers, whether actual or potential or churned, about a product or service (Hennig-Thurau et al., 2004). The influence of eWOM has grown tremendously, as research indicates that user-generated content tends to impact consumer behavior more than traditional marketing efforts (Tsiakali, 2018; Goh et al., 2013).

Given this backdrop, understanding what makes online reviews, or generally, eWOM, be perceived useful is pivotal for marketers. Reviews and comments that are deemed useful add more value to online platforms and help consumers feel more confident about their purchasing decisions (Liu & Park, 2015). So, what makes a review useful? There are several factors to consider, and one important aspect is community engagement.

Many review platforms, like Amazon or 9GAG, allow users to vote on the usefulness of reviews, and comments. These votes not only help separate useful reviews from less helpful ones (Mudambi & Schuff, 2010) but also serve as a quick filtering tool for consumers (Ghose & Ipeirotis, 2008; Liu & Park, 2015). Additionally, the sheer number of reviews can indicate a business's popularity, serving as another guidepost for consumers (Xie et al., 2014). Moreover, studies show that people can assess the credibility of a social media post with the

ratio between the number of likes and comments that it received, as it can show the level of consensus toward the post (Rezaee et al., TBD).

In terms of the content of the reviews and posts themselves, studies have focused on features such as the length and detail of the content. For instance, more elaborate reviews or comments are generally perceived as more useful (Liu & Park, 2015; Castillo et al., 2011). However, the context matters. While detailed posts are valued, longer headlines in online health rumors, for instance, tend to be viewed as less credible (Zhang et al., 2015). The framing of the message in a post also affects its perceived usefulness. According to construal level theory, detailed, specific messages (low-construal) are usually seen as more credible than abstract, generalized messages (high-construal) (Reczek et al., 2018). These detailed reviews can positively influence consumer behavior, especially when they come from credible sources like nano-influencers (Balaji et al., 2021).

Another factor to consider is the use of multimedia elements in reviews and posts. While photos from other customers are generally seen as more credible than those from the company itself (Filieri, 2015), the mere presence of multimedia does not necessarily increase credibility (Li & Sundar, 2022). What matters is how actively people engage with the available multimedia content (Li & Sundar, 2022), and when consumers choose not to engage with the attached media on a post, it does not necessarily make the message appear more credible to them.

5. Dynamics of online reviews and comments

The rise of user-generated content, especially online reviews, has had a profound impact on both consumer behavior and marketing strategies (Tsiakali, 2018; Timoshenko &

Hauser, 2019). While these reviews are highly influential, the sheer volume and diversity of opinions can be overwhelming for consumers. This highlights the importance of understanding the dynamics behind how these reviews aggregate, and how they are affecting consumer's attitudes as a whole.

Accordingly, the area of negative reviews is particularly intriguing. Studies show that inconsistent reviews can damage a brand's reputation and equity (Beneke et al., 2016). However, constructive responses from companies to these reviews can counteract the negative impact and actually enhance their trustworthiness (Könsgen et al., 2018). On the other hand, while you might assume negative comments would hurt a brand, research tells a different story. For example, if you are looking for a restaurant for a special occasion and read mixed reviews, those negative comments might not deter you if you already have a favorable view of the restaurant. Research shows that negative reviews from random users, rather than friends, can increase your purchase intention. This happens because these comments trigger more thoughtful, systematic decision-making as opposed to quick, heuristic judgments. As a result, for those who already have a positive view of the brand, these negative comments can serve as unintentional but effective marketing (Bitter & Grabner-Kräuter, 2016). On the other hand, negative reviews from close friends do not generate the same positive influence; instead, they tend to serve a more diagnostic function, directing attention specifically to the issues mentioned in the comment.

This highlights the concept of "tie strength," or the closeness of the relationship between the content provider and the reader. In the realm of social media, tie strength is also an important factor for advertisers (Wen et al., 2009). While comments from distant acquaintances (or "weak ties") can polarize existing positive attitudes, they can be influential when readers have no prior knowledge or opinions about a brand. In those cases, people tend

to rely on the familiarity of close personal relationships when assessing reviews (Bitter & Grabner-Kräuter, 2016).

When it comes to the emotional tone of a review or its valence, research shows that negative reviews can have more impact than positive ones (Royo-Vela & Casamassima, 2011). However, strategies such as removing negative reviews or disabling the comments section in response to a few negative remarks can be counterproductive. As demonstrated by Rezaee and colleagues (TBD), limiting users' ability to comment on social media posts can actually harm the brand. Specifically, such actions lead to an increase in perceived censorship, which in turn negatively affects users' attitudes toward the brand. More interestingly, as mentioned before, some studies have even found that negative comments can boost sales. The explanation is that negative reviews and comments make people think more carefully about a product, increasing their understanding and confidence, which then translates into purchases (De Maeyer, 2012). Moreover, having a large number of reviews can reduce a potential buyer's sense of risk and uncertainty, further encouraging them to make a purchase (Chen et al., 2004).

6. Attitude toward online content generators

Knowing who is behind an online review or comment matters a lot for its perceived usefulness. Revealing a real identity, especially through a genuine profile picture, can significantly increase the level of trust among users. This idea is rooted in our evolution; historically, most of our interactions were face-to-face, allowing us to gauge trustworthiness through facial cues. Avatars just do not have the same impact on our trust as human faces do (Riedl et al., 2014). However, identity and familiarity and its effect on trustworthiness and perceived credibility is a broader concept than just a mere profile picture.

Generally, different types of familiarity with the content provider can have different effects on the perceived credibility of the information. First, according to identity cue, if the reviewer is someone you personally know or identify with, you are more likely to trust their opinion, as individuals are more receptive to messages presented by members of their own group (Mackie et al., 1990). Second, according to the authenticity cue, expertise matters (Sundar, 2008). A review or recommendation from someone known for their expertise in a particular area carries more weight. For example, health advice from a doctor is generally considered more trustworthy than the same advice coming from a musician (Sundar, 2008). This is why marketers spend so much time finding influencers who mirror their target audience, both from an identity perspective and authenticity (Özbölük & Akdoğan, 2022).

The third type of familiarity relates to reputation cues or celebrity status. If you recognize a reviewer as someone famous or highly regarded, you are likely to find their review more useful (Metzger et al., 2010). Researchers like Ki et al. (2020) have shown that the power of influencers stems from the emotional connections they create, which adds to their credibility. Likewise, Martínez-López et al. (2020) found that when an influencer's persona aligns well with a product, people are more likely to trust both the influencer and the product.

More interestingly, the study points out that a verified account does not guarantee higher perceived credibility. In fact, both verified and non-verified accounts can provide accurate and non-accurate information. Despite this, users tend to see tweets from verified accounts as more trustworthy, but they do not see their provided information as more credible (Vaidya et al., 2019).

Lastly, the number of followers can also serve as a cue for judging an account's credibility. The more followers an account has, the more influential it is seen to be. Interestingly, accounts that follow fewer people are perceived as more independent, further boosting their credibility, thus eventually the ratio between the number of followers to following of an influencer matters in marketing (Valsesia et al., 2020).

7. Attitude toward the sponsored content

When users encounter an action, particularly one that appears to be marketing-related and initiated by another user rather than the company itself, they go through a process to make sense of it. According to Weiner's framework (1974), this involves a three-step process. First, the individual notices a specific behavior. Second, they consider whether this behavior was intentional. Finally, they decide if the person had no choice but to act in that manner due to their circumstances. Applying this to online reviews, if all the reviews consistently agree with each other and the numerical ratings, people are likely to attribute this consensus to the actual performance of the company being reviewed (Könsgen et al., 2018), as it is less likely that all the reviews are just a paid fake comments from the company itself.

There is a controversy in the literature regarding the disclosure of sponsored content to users or not. However, Sah et al. (2018) note that transparency in disclosures can influence user perception. If disclosures are processed automatically by the content provider, they can enhance trust in the influencer's expertise, leading to positive actions like sharing the post or following the promoted brand. However, if users actively ponder over the disclosure, its effects can be lessened or even reversed.

However, the impact of sponsored content extends beyond the moment users first encounter it. For example, brand attitude and purchase intentions can dip when consumers are

directed to a product sales page instead of a generic starting page. This happens because it raises questions about the influencer's impartiality (Stubb & Colliander, 2019). It is important to note that being transparent about sponsored content can improve trust and perceptions of credibility (Vrontis et al., 2021). Yet, the effectiveness of such disclosures can vary based on several factors, such as the style of disclosure and the motives people attribute to the sponsored post.

Additionally, the tone or valence of sponsored content also plays a role in consumer perception. Extremely positive or negative reviews often lead to suspicions of their authenticity (Moon et al., 2019). This is because fake reviews often contain exaggerated sentiments, either highly positive or highly negative, in contrast to more balanced views found in authentic reviews (Ott et al., 2013).

8. Research gaps and opportunities for future research

As explored in preceding chapters, there exists an extensive body of research focused on consumer perceptions of online content. This literature scrutinizes various verbal and non-verbal characteristics inherent to digital media—ranging from text and visuals to interface design—and examines how these elements impact user engagement. Among the metrics evaluated are the level of attention users accord to the content, their perceptions of its credibility and usefulness, and their ultimate attitudes toward the associated brand or their purchase intentions. Despite this wealth of information, there are noticeable gaps in existing research that could have significant practical and theoretical implications (Table 2). These omissions in the literature could represent missed opportunities for both scholars and industry practitioners, as they limit our understanding of consumer behavior in a more externally valid environment.

Table 2
Review of the literature gap on utilizing cues on social media.

Topic	*Subject*	*Central Issue or Research Hypothesis*
General	Cue Interaction in Controlled Experiments	Cues can have various types of relationships, including some cues overriding others, others having a cumulative effect, and some possibly being prioritized by search algorithms (Maheswaran & Chaiken, 1991; Meinert & Krämer, 2022). However, none of the selected articles that explored cues and signals delve into this nuanced interplay. This limitation poses challenges, especially for brands seeking to apply these findings in real-world scenarios.
Attitude Toward Online Contents	Concealing Comments on Social Media Posts	Can hiding comments trigger "reactance" among users (Brehm, 1966), paradoxically making the censored replies more, rather than less, visible? Moreover, can it decrease the attitude toward the brand due to the perceived censorship (Rezaee et al., TBD)?
Dynamics of Social Media Comments	Addressing Negative Posts and Comments	Should one engage with negative posts or comments on social media, considering that responding may increase their visibility? Is it advisable to be selectively responsive to negativity on social media? What might be the most effective tone or strategy for responding to complaints on these platforms?
Attitude Toward Online Content Generators	Introduction of Paid Verification Badges	Vaidya et al. (2019) showed that verified accounts were generally seen as more trustworthy. However, recent changes on platforms like Twitter and Instagram have introduced the option to purchase these verification badges. This development raises concerns about the ongoing credibility of Vaidya et al.'s findings. In particular, does the availability of purchasing a verification badge erode its previous associations with reputational validity?
	Crafting an Effective Bio	Balancing precision and comprehensiveness in a social media bio is a common concern. While research suggests that using emojis can reduce perceived competency (Vareberg et al., 2023), it is also known that they can convey additional information and enhance text comprehension, particularly within the constrained space of a bio (Boutet et al., 2021). The question arises: should emojis be included in a bio?

Topic	*Subject*	*Central Issue or Research Hypothesis*
Attitude Toward the Sponsored Contents	Contrasting Mixed and Homogeneous Advertisements	De Maeyer (2012) discovered that negativity can stimulate systematic thinking about an advertised product, improving comprehension and potentially leading to more purchases. Moreover, the systematic thinking can have longer attitude change than peripheral one (Petty, 2018). Accordingly, the perception of impartiality in user-generated reviews or sponsored content is crucial for their effectiveness and influence on buying choices (Stubb & Colliander, 2019). Given these findings, it is essential to explore the potential advantages of incorporating negative aspects in the sponsored content.
	Identifying the Optimal Timing for Reciprocity	For YouTube sponsorship, Is it more beneficial to initially provide the desired content and then introduce the sponsored segment? Could adopting a "give before you ask" strategy increase the viewer's willingness to engage with the sponsored message (Regan, 1971)? Alternatively, is it more effective to do it at the outset when the impact is greatest?

Firstly, as previously mentioned, there are intricate dynamics between central and peripheral cues in shaping user perception and behavior. These cues can have varying types of relationships: some may override others, some could have a cumulative effect, and some may even be prioritized by searching algorithms (Maheswaran & Chaiken, 1991; Meinert & Krämer, 2022). However, none of the selected articles that discussed cues and signals focused on this nuanced interplay. In the majority of experiments conducted for these studies, all confounding variables were carefully controlled, and manipulations focused solely on the cue under investigation. While such an approach is invaluable for understanding the isolated impact of individual cues, it falls short of offering a comprehensive view. This is particularly problematic for brands looking to apply these findings in real-world scenarios. Decision-making is a complex process that involves a host of variables, and it cannot be fully understood by examining cues in isolation. The multifaceted "black box" of decision-making takes into account both central and peripheral cues, often in ways that research has not yet explored. For instance, the recognition heuristic is often described as non-compensatory,

implying that once this cue is received, individuals generally cease to seek additional cues for decision-making (Goldstein & Gigerenzer, 2002), therefore the experimental results regarding this cue can have high external validity with the real-life scenario. However, it is important to note that this characteristic may not universally apply to other cues discussed in this article. Moreover, it is also crucial to acknowledge that there is no assurance that individuals will encounter this recognition cue when presented with a variety of other cues and signals in a given post. This adds another layer of complexity to the understanding of how cues are utilized and prioritized in the decision-making process.

Therefore, it is important to understand how a proposed cue operates in conjunction with other prominent cues that may influence user decision-making. For instance, if the follower-to-following ratio is a key metric in assessing an account's credibility (Valsesia et al., 2020), what happens to that assessment when other credibility cues are introduced? How does the presence or absence of a verification badge (Vaidya et al., 2019), or the account holder's demonstrated level of expertise in a relevant field (Sundar, 2008), affect user perception? What about when we see lots of negative comments in a post from a person with a perfect follower-to-following ratio? Which one is the most important? Which one can override the effect of others without any problem? In summary, understanding the overarching algorithm or cognitive framework that users employ to weigh various established cues could have a profound impact on both theoretical inquiry and practical application. Without this broader understanding, recommendations based solely on isolated cues may offer limited utility for brands aiming to influence user behavior in a meaningful way.

Secondly, Rezaee and colleagues demonstrated that restricting users' ability to comment on a post can negatively impact attitudes toward a brand by increasing the perception of censorship (TBD). However, a related yet distinct phenomenon exists on social media platforms such as Twitter: the practice of "hiding comments." Unlike restricting the

ability to comment, hiding comments does not stop the negative comments altogether but rather minimizes their visibility. Interestingly, this action comes with a visual cue, letting users know that certain comments have been hidden. This raises a pertinent question: does the act of hiding comments generate a form of "reactance" among users toward the censored replies, thereby making those comments more visible rather than less (Brehm, 1966)? In other words, could the act of hiding comments pique users' curiosity to such an extent that they end up reading the hidden comments even more than they would the visible ones?

Thirdly, research indicates that company responses to negative reviews can help mitigate their adverse effects on perceived trustworthiness (Könsgen et al., 2018). However, the dynamics of social media platforms introduce a unique factor: visibility.

This point highlights the distinctive characteristics found in social media platforms when compared to platforms primarily focused on reviews, like Amazon's review section. For instance, responding to a negative comment on a platform like Twitter, amplifies its visibility to other users, when it is not the case for review platforms. When a company replies to a complaint tweet, that interaction becomes more visible to those who engage with the company online. This nuance may necessitate a reevaluation of Könsgen's recommendations.

Furthermore, existing literature lacks studies that specifically explore best practices for responding to complaints and negative comments on social media. For example, should the response be positive in tone? Does the use of emojis and friendly language enhance communication, or does it compromise the message's authenticity? Answering these questions is crucial for PR teams who are continually navigating negative comments and controversies on social media, often from dissatisfied or churned consumers.

Another area for investigation involves the impact of verification badges on social media credibility. Vaidya et al. (2019) demonstrated that accounts with verified badges were generally perceived as more trustworthy. However, recent changes on platforms like Twitter

and Instagram now allow users to purchase these verification badges. This development raises questions about the continued validity of Vaidya et al.'s findings. Specifically, does the ability to buy a verification badge undermine its previous connotations of reputational validity? Furthermore, could this new dynamic potentially backfire, causing users to view badge holders more skeptically due to the perception of purchased privilege?

Another significant gap in existing research pertains to the impact of a user's social media "bio" on their perceived credibility and authenticity. While studies have shown that extreme tones in reviews can arouse suspicion (Moon et al., 2019), it remains unclear how these principles translate to the limited space of a social media bio, where users often have to encapsulate their identity in just a few words, which may lead them to use extreme and concise wording. Additionally, the effect of specific elements within a bio, such as emojis or academic and professional titles (e.g., Dr., MBA), has not been extensively studied. Do these elements enhance a bio's credibility or diminish it? For example, we know that using emojis can decrease perceived competency (Vareberg et al., 2023), but on the other hand, we also know that they can convey more information and increase the information procession of the text (Boutet et al., 2021), especially in the limited space of a bio. Similarly, does the inclusion of formal titles like "Dr." or "MBA" in a bio automatically confer a sense of authority and credibility, or could it induce skepticism in certain contexts? Answering these nuanced questions could offer valuable insights for individuals looking to effectively manage their personal branding on social media. Understanding the dynamics of bio construction could also assist companies in creating more authentic and relatable online profiles, ultimately influencing public perception and trust.

Furthermore, existing research provides intriguing insights into how negative comments can sometimes positively influence sales, by increasing the attention to them. For example, De Maeyer (2012) found that negative comments can prompt people to think more

systematically about an advertised product, thereby increasing their understanding and ultimately leading to more purchases. Furthermore, it is worth noting that altering peripheral thinking to systematic thinking can influence the duration of the effect on attitude change (Petty, 2018). On the flip side, the perception of impartiality in user-generated reviews or sponsored content also plays a critical role in their effectiveness and influence on purchasing decisions (Stubb & Colliander, 2019). Given these insights, it becomes important to investigate the potential benefits of featuring mixed sponsored content—that is, content that highlights both the positive and negative aspects of an advertised product or service. This approach could offer dual advantages. First, it could enhance the perceived impartiality of the reviewer, making the review or content appear more trustworthy. Second, presenting negative points could engage users in more systematic thinking, thereby making the advertisement more memorable and effective (De Maeyer, 2012). For instance, should companies consider strategically highlighting their product's objective strengths while also openly acknowledging its subjective shortcomings? Such an approach could potentially make their marketing efforts more impactful, as it leverages both the power of impartiality and the cognitive engagement triggered by a balanced perspective.

Finally, existing research has explored how the disclosure of sponsored content can influence users' perceptions of both the advertisement and the reviewer (Sah et al., 2018). However, the literature is notably silent on the optimal timing for such disclosures. For example, YouTube influencers often embed sponsored segments within their videos, but it remains an open question as to when is the most effective moment to introduce these paid endorsements. Should they appear at the beginning, mid-point, or end of the video?

One critical factor that could influence this timing is the principle of reciprocity. When viewers engage with content from their preferred influencers, they are, in essence, receiving something they value. This raises the question: Would it be more advantageous to first deliver

the sought-after content and only then dive into the sponsored segment? Could a "give before you ask" strategy enhance the viewer's receptivity to the sponsored message (Regan, 1971)?

Furthermore, Burger et al. (1997) noted that the urge to reciprocate a favor diminishes as the time between the initial favor (in this case, the valuable content) and the opportunity for reciprocation (such as liking, subscribing, or visiting the website of the sponsored content) expands. This introduces a complex trade-off. Presenting sponsored content at the video's start would ensure maximum visibility since all viewers clicking on the video would see it. However, introducing sponsored content after delivering valuable information could heighten viewers' sense of obligation to reciprocate, potentially leading to higher engagement rates. Therefore, is it more effective to place the sponsored segment immediately following the video's peak to maximize the likelihood of reciprocation, as suggested by Burger et al. (1997)? Alternatively, Would it be more effective to slightly temper the immediate urge for reciprocity and introduce the sponsored segment at the end of the video? This approach could potentially increase the likelihood of viewers following through with a call-to-action, such as making a purchase or subscribing, as viewers are generally less inclined to interrupt their viewing experience to switch platforms for making a purchase. This question presents an important dilemma and highlights an area for further study, as a nuanced understanding of these dynamics could offer invaluable insights for content creators and advertisers alike.

9. Conclusion

This essay underscores the significance of paying attention to both verbal and non-verbal cues and signals (Paivio, 2013) that users commonly employ in their everyday social media interactions. This emphasis arises from the context of information overload and a reduced inclination or motivation for engaging in systematic thinking (Pee, 2012).

Consequently, this paper is dedicated to the review and categorization of articles, primarily published between 2013 and 2023 in journals related to marketing, business, and communication. Ultimately, it aims to draw attention to critical research gaps that have the potential for substantial contributions to both academia and practical applications, particularly in the realm of understanding the attitudes towards online content and content providers in social media.

In conclusion, it becomes clear that marketers must remain exceptionally mindful of the cues and signals within their selected digital communication channels. Even seemingly minor cues can wield diverse levels of influence on their target audiences, and any attempts to counteract these cues may, in fact, produce unintended adverse effects. Therefore, comprehending the underlying dynamics of these cues emerges as a crucial imperative. Furthermore, it is crucial for marketers to understand that, before implementing any recommendations based on academic findings, conducting tailored tests within the unique context of their own business is equally vital when formulating an effective social media communication strategy, as the academic research may have limitations, often stemming from controlled experimental settings.

Getting Ratioed: The Effect of the Comment-to-Like Ratio on the Perceived Credibility of Social Media Posts

ABSTRACT

This article examines a social media phenomenon called 'ratioing,' which describes a high ratio between the number of comments and the number of likes that a social media post such as a tweet receives. Specifically, we develop and test theoretical arguments that consumers perceive posts with a high ratio to be less credible than posts with a low ratio. A combination of one field study and five experiments provides evidence for this proposition. The studies also show that the reason for this effect is that consumers perceive higher ratios as an indication that others agree less (i.e., low degree of consensus) with the content of those posts compared to posts with lower ratios. An incentive-compatible study using social media posts about new products demonstrates that the ratio also affects consumers' purchase behavior. This article contributes to our theoretical understanding of how consumers process information on social media and provides practical insights for companies that use social media as a means of communication for their products and brands.

Keywords: Social Media Indicators, Perceived Credibility, Degree of Consensus, Information Processing on Social Media.

1. Introduction

Consumers constantly process information from various sources. The one source that has become more important in recent years is social media (Bradley, 2010). A key difference to traditional media is that companies use them not only to communicate information about their brands and products to consumers, but also to interact with consumers. Moreover, social media enable consumers to directly express their opinions for instance, via posting content, commenting on already posted content, and liking that comment.

Not surprisingly, research shows that while consumers use similar characteristics (e.g., by the source, the familiarity, or the design of a website) to judge information from social media as for traditional media (Kalbfleisch, 2003; Lucassen & Schraagen, 2013; Wells et al., 2011), they also form their evaluations based on different characteristics, such as the profile picture of the source (Xu, 2014) or the grammar and punctuation of a post (Morris et al., 2012).

One particularly novel and unique aspect of social media is the *social* dimension; users can typically see the reactions and responses of fellow users in the form of comments, shares, and likes. Despite the common focus of practitioners on, at least some, of those indicators, research has to date dedicated only limited attention and effort to examining the impact of these social indicators on consumers' evaluation of social media posts.

The key proposition of this paper is that consumers use the number of comments relative to the number of likes that a social media post receives (hereafter called the 'ratio') to assess the credibility of that post. Imagine reading two tweets about a new technology product on Twitter. One tweet (tweet A) has 1000 comments and 10 likes. Another tweet from the same company (tweet B) has 10 comments and 1000 likes. Which of those two tweets, tweet A or tweet B, would have a greater effect on consumers' purchase decisions?

In this paper, we propose that the former post (tweet A), which the social media community describes as 'being ratioed' (Minot et al., 2021) is perceived to be less credible than the latter post (tweet B) because consumers use the ratio between the number of comments and likes as a cue of the perceived degree of consensus of fellow consumers with the post's content. In other words, the lower ratio is seen as community consent, and the higher ratio as community dissent.

The insights of one field study and five experiments support our proposition that consumers use the ratio of a post as a cue to judge that post's credibility. The field study provides preliminary evidence for the process by which we hypothesize the ratio affects perceived credibility. Study 1 provides controlled evidence for the main effect, of the ratio on perceived credibility. Study 2 demonstrates that the ratio (and not comments or likes alone) cause of perceived credibility. While Study 2 also provides correlational evidence (using mediations) that the ratio affects perceived credibility via perceived degree of consensus, Study 3 demonstrates the causality of this process. Study 4 demonstrates the downstream consequences of the ratio using an incentive-compatible design and consumer choice as the dependent variable. Finally, Study 5 reveals a boundary condition by demonstrating that social media posts that ask for feedback do not lead to reductions in perceived credibility. Thereby, this study gives an idea of how to reduce the adverse effects of high comment-to-like ratios, which might naturally evolve for controversial topics.

The paper proceeds as follows. First, we review the literature on consumer information processing on social media to develop a theory on how the comment-to-like ratio affects the perceived credibility of a post and eventually the consumer purchase decisions. Second, we present the results of our six studies designed to test the theory and rule out alternative explanations. Finally, we conclude with a general discussion of the implications of our findings for researchers and practitioners.

2. Consumer information processing on social media

Humans' cognitive system is known to have limited processing capacities, and when subjected to an overload of information, the quality of decision-making is likely to be adversely affected (Gross, 1964). The emergence of social media has made consumers even more likely to experience an information overload (Rodriguez et al., 2014). As a result, consumers face a trade-off between efficiency and accuracy when they judge the credibility of information (Fogg, 2003). Consumers often respond to that trade-off by using specific cues and heuristics that help them to make decisions (Tversky & Kahneman, 1974) and judge social media posts (Hilligoss & Rieh, 2008; Ranganathan, 2012).

Existing research on social media revealed that people use cues—mainly contextual or behavioral—to judge the credibility of posts. Firstly, contextual cues, such as a proper profile picture and username, having a verified badge, and the number of followers can increase perceptions of authenticity (Morris et al., 2012; Vaidya et al., 2019). Furthermore, the use of non-standard grammar and punctuation mistakes can reduce the perceived credibility of a social media post (Morris et al., 2012). More importantly, the number of likes and shares can affect the perceived credibility as they send signals about the importance, relevance, and reliability of information (Avram et al., 2020). Finally, the perceived level of effort that the author put into the post could affect the perception of credibility. For instance, using supporting multimedia, such as pictures and video (Kiousis, 2006), and the length of the post (Castillo et al., 2011) have been shown to increase the perception of credibility.

Secondly, social cues about the source, such as identity cues and authority cues, can play a significant role in information processing on social media (Lin, 2016). Identity cues include the degree to which consumers base their judgment of a post based on the identity of the individual or organization that posts that tweet (Sundar, 2008). For instance, people trust

information from a known source (e.g., a friend) more than from an unknown source (i.e., a stranger). Authority cues describe the extent to which a person or organization with authority posts information and thus relates to the source (Sundar, 2008). For example, consumers tend to assess information about a vaccine as more credible when a doctor (vs. a musician) posts that information given the former's expertise.

Finally, Valsesia et al. (2020) showed that consumers use the number of accounts an individual or organization follows and is being followed by to judge the individual's or organization's perceived influence and perceived autonomy. In particular, the more users follow an account, the more it is perceived as influential, and the fewer other users an account is following, the more it is perceived as autonomous, which eventually affects consumers' perceptions of the credibility of that account (i.e., individual or organization). More importantly, Valsesia et al. (2020) provide initial evidence that consumers tend to process information on social media in the form of ratios. When consuming information on social media (e.g., reading posts of the accounts that they follow), consumers often lack the motivation and/or ability to thoroughly evaluate (e.g., by fact-checking the information) that information. Therefore, consumers tend to use heuristics to judge the credibility of that information, for example by using the number of followers on Twitter (Lee & Sundar, 2012).

This prior work has shed some initial light on several contextual and social cues regarding the source and the posts on social media that affect the perceived credibility of a post. In this paper, we follow the latter avenue and examine a so-far overlooked social cue by which consumers judge the credibility of information on social media, namely the effect of 'being ratioed.' The key proposition of this paper is that consumers use the number of comments relative to the number of likes (i.e., the ratio) that a post receives from fellow consumers to judge the credibility of a social media post. In the next section, we develop a theory on consumers' interpretation of the comment-to-like ratio of social media posts and

provide substantive arguments for why the ratio, and not the individual indicators alone, serves as the basis for consumers' credibility judgments of those posts.

3. Psychological consequences of the comment-to-like ratio

The comment-to-like ratio consists of two parts, the number of comments and the number of likes. It is important to understand the individual parts to understand the consequences of the ratio between them.

Making a comment on (i.e., responding to) a social media post allows people to add additional context and nuances to the initial post but also their own opinions about the content of the original post. Social media comments allow users not only to share their own views but also to observe the overall responses of other viewers (Waddell & Sundar, 2017). In addition, via a comments section, users can react to the news, whether it is political, technical, or a piece of gossip from a friend (Almgren & Olsson, 2016). Comments may not only signal agreement with a social media post but can also express disagreement to some extent (West, 2015). Nevertheless, making a comment also requires people to dedicate some amount of effort to crafting the response.

By contrast, liking a social media post merely requires a click. More importantly, consumers can interpret 'liking a post' in various, however, more positive ways. By liking a post people can show others that they have seen the content, or, that they agree with the content of the post (Levordashka et al., 2016). Muntinga and colleagues (2011) suggest that the ability to easily consent to a social media post by simply clicking the like-button makes users who express agreement with a social media post through commenting oftentimes provide a like, too—in addition to their comment.

In sum, when considering comments and likes on a social media post there are four options, three of which are relevant for the ratio: Commenting, liking, or both (i.e., commenting and liking). While likes predominantly signal agreement, and providing both likes and comments signals strong agreement, we argue that only commenting should be interpreted as a disagreement or, at least, not full agreement.

While the number of likes alone could increase the perceived credibility of a post to some extent (Avram et al., 2020), we argue that the ratio between the number of comments and likes has a significantly greater influence on perceived credibility. Consistent with prior work on consumer information processing on social media (Valsesia et al., 2020), neither the number of comments nor the number of likes alone, but the ratio between both should affect the perceived credibility of a social media post.

Using ratios between two indicators to form a judgment is not new to social media. Valsesia et al. (2020) showed that consumers use the ratio between the number of accounts an individual or organization follows and is being followed to judge the individual's or organization's credibility. Furthermore, De Vries (2019) posited that the ratio between the number of likes and the number of followers can influence the perceived credibility of the account owner. Likewise, Permana and Meinarni (2021) have demonstrated the impact of certain parameters on TikTok, such as the 'following-to-like ratio' and 'video-share-to-video-comment ratio,' on the perceived credibility of TikTok accounts. The use of ratios extends beyond TikTok, with Mohan et al. (2018) finding that consumers assess the fairness of salaries within a company by comparing the CEO's compensation to that of the median employee, rather than evaluating the CEO's salary in isolation.

Given this work and our argumentation that consumers understand making a comment as disagreement with the content while liking shows more overt agreement (or, consensus), the comment-to-like ratio should be inversely related to consumers' perception of the level of

agreement among their peers with the post. Specifically, consumers should perceive that their peers have a lower degree of consensus with the content of a social media post the more comments (vs. likes) that post receives, while the opposite is true for a post that receives relatively fewer comments than likes.

Perceived consensus can be defined as one's perception of the opinion of the majority of a group of people about an idea or behavior (Chaiken, 1987). In other words, perceived consensus refers to the views of the majority versus a minority of people and not the number of people who believe/do something. Importantly, perceived consensus reflects a ratio as it expresses an individual's perceptions of the proportion of people who agree (vs. disagree) with a statement, or as Todorov et al. (2002) stated, "the presentation of the result of an opinion poll in which the majority of respondents agree with the advocate position."

The importance of perceived consensus in shaping attitudes and message credibility has been highlighted in the literature (Chaiken & Stangor, 1987). This is particularly pertinent in the online context, where perceived consensus can be more visible and influential in shaping attitudes towards e-commerce messages (Park and Lee, 2008). Furthermore, the level of perceived consensus can mediate the relationship between the number of likes and individuals' attitudes toward a particular topic, as demonstrated in the context of advertised products (Lee, 2015). Notably, Perceived consensus can be either implicit, which is based on the majority opinion, or explicit, which refers to the percentage of people who agree with a statement (Erb et al., 2006).

Recall the example stated in the introduction. Tweet A received 1,000 comments and 10 likes and Tweet B received 10 comments and 1,000 likes. While the former has a ratio of 100/1, the latter has a ratio of 1/100. We argue that consumers perceive the consensus with tweet A (ratio: 100/1) to be lower than for Tweet B (ratio: 1/100). Eventually, they should perceive Tweet A also to be less credible than Tweet B. Thus, we hypothesize:

H1: The greater the comment-to-like ratio of a social media post, the lower the perceived credibility of that post.

H2: The effect of the comment-to-like ratio of a social media post on perceived credibility is mediated by the perceived degree of consensus of fellow consumers with that post.

As we hypothesized, an imbalanced comment-to-like ratio often indicates a significant number of people who engage with a post by commenting without also liking it. This can serve as an indicator of the collective sentiment among individuals who have interacted with the post, similar to a large-scale peer review process within the social media sphere. However, it is important to note that commenting does not always signify agreement or disagreement; it can also be a simple reaction to the content of the post itself. Consequently, altering the post's content in a way that guides user interpretation away from mere disagreement may mitigate the negative impact of a skewed comment-to-like ratio.

In this paper, we propose that explicitly seeking feedback within a post can transform the way comments on ratioed posts are perceived, emphasizing solicited feedback rather than outright negative disagreement. This shift acknowledges that individuals may provide feedback without necessarily liking a post.

H3: The effect of the comment-to-like ratio of a social media post on perceived credibility is moderated by the feedback solicitation in the post.

4. Overview of studies

We conducted a total of six studies (one field study, and five experiments) to test our two hypotheses and examine the effect of the comment-to-like ratio on the perceived credibility of the social media posts. The field study uses randomly collected social media

posts (N = 1,400) to provide initial evidence that the comment-to-like ratio is correlated with the degree of consensus with a social media post. Study 1 (N = 604) tests the main hypothesis of this paper (H1) in a controlled experiment. The results show that a high ratio decreases the perceived credibility of a social media post while a low ratio increases the perceived credibility of that post. Study 2 (N = 1,002) is a pre-registered between-subjects experiment that uses a more conservative design. The results not only corroborate that a high (vs. low) ratio decreases (increases) the perceived credibility of a social media post but also shed light on the process underlying this effect by revealing degree of consensus as a key mediator. Study 3 (N = 403) is pre-registered and manipulates the degree of consensus and thus provides *causal* evidence on its role as the mediator of the effect of the ratio on perceived credibility and thus support for H2. Study 4 (N = 1,000) uses a consequential setting with high realism to show that the ratio not only affects perceived credibility but also consumers' actual purchase behavior. The pre-registered Study 5 (N = 401) demonstrates that soliciting feedback in a post can attenuate the negative effect of the ratio.

Together, the studies provide strong evidence that a high ratio reduces the perceived credibility of a social media post and that the (low) perceived degree of consensus causes the effect. According to the studies, the comment-to-like ratio of the posts on social media can have an impact on consumer purchase intention and behavior. However, firms can use strategies to frame the content of the post in a way that bypasses the effect of ratioing on perceived credibility, particularly for potentially controversial posts that are likely to receive a lot of comments.

5. Field study

The objective of the field study is to provide an initial test of the relationship between the ratio of a social media post and consumers' perceptions of the degree of consensus with the content of a social media post. This study examines the main proposition of this paper in a naturalistic setting (with high ecological validity).

5.1. Procedure

For this study, we randomly collected 1,400 social media posts from five consecutive days in January 2021 from Twitter. We obtained the number of comments and the number of likes that each of the social media posts received. We also collected the first ten comments shown in response to the post. In the second phase of this study, we employ a decision tree analysis to assess the relative significance of the comment-to-like ratio when compared to other coded variables.

5.2. Measurement

5.2.1. Ratio

The ratio was calculated by dividing the number of comments on each post by the number of likes.

5.2.2. Degree of consensus

Two research assistants blind to the hypotheses of this project (and the ratios of the tweets) coded each of the comments regarding the degree to which the responses agreed with the social media post on a 9-point scale (1 = "No agreement" to 9 = "High agreement").

5.2.3. Content

We enlisted the help of a third research assistant to label the content of the tweets. Their task was to read each tweet and categorize it as 'news', 'daily tweets' (such as talking

about daily life and personal experiences), ‘ads’, or ‘response’. This approach enabled us to ensure that the content of the tweets did not confound the relationship between the degree of consensus and the ratio. Appendix 11 provides a summary of the descriptive statistics for the field study and the subsequent experiments.

5.3. Results

A linear regression model with the ratio as the independent variable and the degree of consensus as the dependent variable showed that the ratio was marginally correlated with the degree of consensus (B = -.03, t = -1.75, p = .08). We repeated the linear regression but controlled for the content type, word count, number of used emoji, and number of used hashtags. The results showed that the ratio was significantly negatively correlated with the degree of consensus (B = -.03, t = -2.05, p = .04). We again repeated the linear regression and included the number of comments and likes as additional control variables (besides the previously reported ones). The results of the prior regression prove to be robust given that the effect of the ratio on the degree of consensus remained significant and negative (B = -.03, t = -2.04, p = .04). Neither the number of comments (B < .001, t = .19, p = .85) nor the number of likes (B < .001, t = -.88, p = .38) alone had a significant effect on the perceived degree of consensus. Table 3 shows the results of each of the three regressions.

Table 3
Results of the regression analyses (Field Study).

	Non-controlled regression				*Control: Tweet characteristics*				*Controls: Number of comments & likes*			
	B	**SE**	**t**	***p***	**B**	**SE**	**t**	***p***	**B**	**SE**	**t**	***p***
Ratio	-.03	.02	-1.75	0.08	-.03	.02	-2.05	.04	-.03	.02	-2.04	.04
# of likes	-	-	-	-	-	-	-	-	<.001	<.001	-.88	.38
# of comments	-	-	-	-	-	-	-	-	<.001	<.001	.18	.85
Type: Daily tweets	-	-	-	-	-.07	.28	-.25	.80	-.07	.29	-.24	.81
Type: News	-	-	-	-	-1.22	.29	-4.21	<.001	-1.22	.29	-4.20	<.001
Type: Response	-	-	-	-	-.37	.30	-1.22	.22	-.37	.30	-1.22	.22
# of emojis	-	-	-	-	.05	.03	1.66	.10	.05	.03	1.66	.10
# of hashtags	-	-	-	-	.09	.06	1.48	.14	.09	.06	1.45	.15
# of words	-	-	-	-	<.001	.003	.19	.84	<.001	.003	.17	.89
Constant	6.39	.05	138.65	<.001	6.83	.29	23.47	<.001	6.83	.29	23.48	<.001
R^2			.001				.10				.10	

The ratio was negatively correlated to the degree of consensus (as coded by the assistants), while the number of comments or likes was uncorrelated to that measure. The correlation between the ratio and the degree of consensus supports the proximal path of the proposed mediation (H2).

5.4. Decision tree analysis

We additionally conducted a decision tree analysis to ascertain the hierarchical importance of the comment-to-like ratio in relation to other coded variables, including the number of likes, comments, retweets, emojis, and more. Furthermore, we employed three dictionaries, namely LIWC, PARA, and Harvard, to code various aspects of the tweets that might influence perceived credibility.

The decision tree analysis (Figure 1) reveals that, following the content type (whether the tweet falls into categories such as advertising, news, daily talk, or is a response to another tweet), the comment-to-like ratio emerges as the most crucial predictor for determining the level of agreement between comments on a tweet and the original tweet. Notably, it surpasses the significance of the absolute counts of likes and comments, further reinforcing our hypothesis (Additional details on the decision tree analysis can be found in Appendix 12).

Importantly, the results of the decision tree reveal a critical threshold of .24 for the ratio, signifying its substantial impact on the level of agreement. In the dataset comprising random English tweets collected in July 2023 (N = 45,435), this threshold corresponds to approximately 15.2% of the total tweets, highlighting its significance in the real data.

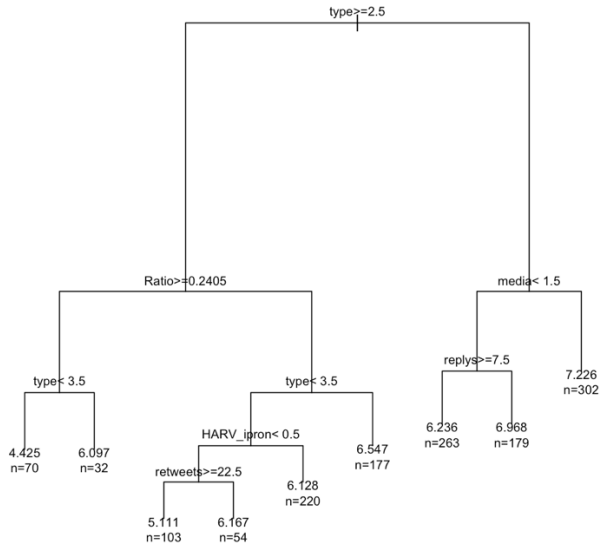


Fig. 1. Hierarchical Predictors of Level of Agreement (Depth Limited to 8 Levels)

5.5. Discussion

This field study provided a first indication that the ratio and the degree of consensus are negatively correlated. In particular, the greater the comment-to-like ratio of a social media post, the lower the degree of consensus with the content of the original post. Furthermore, the decision tree underscores the relative importance of the ratio as a predictor of the level of agreement with the tweet, underscoring the substantial contribution of our findings.

As any field study, this one has its limitations: This study used uncontrolled field data with a correlational design and thus cannot show causality. Our measure of perceived consensus builds on the last ten comments on each tweet. Using the last ten comments also could affect our result as the last ten comments are not shown at random (i.e., they depend on their popularity as well as information about the account). However, this is also the case in real life as customers do not always see and read all comments on a social media post. Finally, this study provided correlational evidence on the effect of the ratio on the mediator—degree of consensus. However, this study did not examine the key hypothesis of this article, i.e., that the ratio negatively affects the perceived credibility of a tweet.

6. Study 1

The objective of Study 1 is to examine whether a high (vs. low) ratio decreases the perceived credibility of a social media post (H1). It uses an experimental design to manipulate the ratio.

6.1. Procedure

We recruited 604 participants (M_{Age} = 40.7, SD_{Age} = 13.1, 43.6% male) from Amazon Mechanical Turk ("MTurk") in exchange for a monetary payment. Using an online sample fits the research question well, given that the phenomenon of 'ratioing' is inherently observed online (i.e., on social media).

This study used a mixed-subjects design with three conditions (ratio: high vs. low vs. control condition), whereby each participant saw and rated a total of ten social media posts of a total of twenty social media posts. The posts covered various fictitious topics, such as news about non-existent events or advertisements for non-existent products, to prevent any prior knowledge that participants may have had about the topics from affecting the results. After introducing participants to the survey, participants were randomly assigned to view the ten posts in a sequence (one by one), with a varying number of comments and likes for each post shown below it. Thus, each person saw different posts of all conditions (e.g., the first post could be in the high ratio condition, while the second post could be in the low ratio condition). This sequential exposure resembles consumers' real experience on social media where different posts have different ratios.

The ratio was manipulated as follows: In the 'high ratio' condition, the ratio between comments and likes was approximately 440 (i.e., 440 comments and 1 like, or 865 comments and 2 likes, etc.), while in the 'low ratio' condition, the ratio was approximately 1/440 (i.e., 1 comment and 440 likes, or 2 comments and 865 likes, etc.). The ratios varied within the conditions (440± 5%, and 1/440± 5%) to avoid participants guessing the hypothesis of this

work. These subtle variations also allowed us to control for the number of likes and comments of each post in a subsequent regression analysis. In the 'control' condition, participants neither learned about the number of comments nor likes on a post. An example of the stimuli is available in Web Appendix 1.

Importantly, the stated manipulations on the ratio (number of comments and likes) were based on 70,000 posts that we randomly selected and obtained from Twitter. The mean comment-to-like ratio for this dataset was 0.01, while the median was 0. To create social media posts with a low ratio, we used the ratios from the bottom 5th percentile, which were 440, and for high ratios, we used the reciprocal of this value, 1/440. While this manipulation seems heavy-handed, future studies employed more subtle manipulations.

6.2. Measurement

At the initiation of the study and following the attention check, we measured participants' levels of attention to textual and non-textual information, as well as their susceptibility to social influence, aiming to control for potential confounding variables in our main model. This was accomplished using three one-item Likert-type scales, wherein participants were prompted to indicate the extent to which they focus on textual information within the provided content, the extent to which they direct their attention to non-textual information in the text, the extent to which display susceptibility to social influence when making decisions regarding the information. Ratings were gathered on three one-item Likert-type scale, ranging from 1 (indicating "Totally disagree") to 7 (indicating "Totally agree").

Participants evaluated the perceived credibility (M = 4.49, SD = 1.59) using a modified version of the 9-item scale of Ad Trust (Soh et al., 2009). Specifically, participants rated "This post was…" "honest," "truthful," "credible," "reliable," "dependable," "accurate," "factual," "complete," and "clear" ranging from (1 = "Strongly Disagree" to 7 = "Strongly Agree;" Cronbach a = .96).

Participants were tasked with assessing multiple aspects of the tweets in order to control for potential variables that could impact their perceived credibility. These aspects included quantifying the order of magnitude of likes and comments for each tweet, evaluating the perceived sentiment of the tweets, showing the level of interest toward the content, and analyzing the length of the tweets. These measurements were conducted with the primary aim of exploration. All the aforementioned variables were evaluated using single-item Likert-type scales.

Before asking demographic questions, participants responded to two hypothesis-guess questions. Participants could guess the hypothesis in an open-ended format. Participants also could choose one of six available answers regarding what they believed was the purpose of the study. All available answers were reasonable, such as "the impact of post grammar on perceived credibility" or "the influence of the number of likes on perceived credibility." Finally, all participants indicated their age (in groups), gender, and experience in using social media.

6.3. Results

6.3.1. Perceived credibility

Given that the responses are nested due to the within-subjects design, we used a linear mixed model to analyze the data, controlling for tweets, participants, number of comments, number of likes, and the content type. In addition, we added a dummy variable to distinguish tweets about news from commercial topics. The results of the linear mixed model showed that a high ratio (marginally significant) reduces perceived credibility (B = -.14, SE = .07, t = -1.94, p = .052). Thus, a high ratio reduces perceived credibility, supporting our key proposition. More interestingly, neither the number of comments nor likes had a significant effect on perceived credibility. The control condition, which did not have any number of comments or likes, was excluded from the first regression as no information on the number of

likes or comments was available. Furthermore, we conducted an additional iteration of the model, this time incorporating all of our control variables. These variables encompassed factors like focus on textual and non-textual content, susceptibility to social influence, and the level of experience in utilizing social media platforms. The results show a significant effect between our manipulation and perceived credibility (B = -.14, SE = .07, t = -1.97, p = .048).

We used a second linear mixed model, only with controlling the tweet, its' content type, and the participants, to be able to include the control condition which did not have any comments or likes. The results showed that respondents perceived posts with a high ratio as less credible than posts with a low ratio (B = -.23, SE = .03, t = -7.56, $p < .001$) and that also less credible than in the control condition (B = -.15, SE = .03, t = -5.08, $p < .001$). The control condition was not significantly different from the low ratio condition (see Figure 2). Thus, a high ratio decreased the perceived credibility of the social media posts.

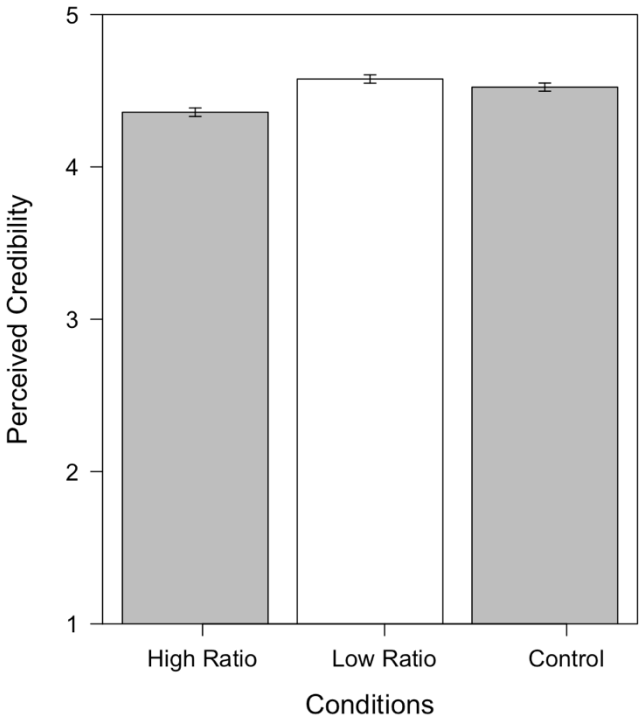


Fig. 2. A High (vs. Low) Ratio Reduces (Increases) the Perceived Credibility of a Social Media Post.

Additional irritation of this model Including all the measured control variables, such as focus on textual and non-textual content, susceptibility to social influence, the level of experience in using social media, perceived length, perceived interest toward the content, and perceived sentiment, replicates the significant difference between the High Ratio group and Low Ratio (B = -.19, SE = .03, t = -6.58, $p < .001$). However now, the difference between the control group and the Low ratio became marginally significant (B = -.05, SE = .03, t = -1.69, $p = .09$).

6.3.2. Content type as moderator

We conducted another linear mixed model regression with ratio as the independent variable, and content type as the moderator while controlling for tweets and participants as fixed effects. The results indicated that content type did not have a significant interaction effect (t = -.36, $p = .72$, and t = -.49, $p = .62$), while the main effect again only replicated for the High and Low Ratio group (B = -.22, SE = .04, t = -5.10, $p < .001$) and not for the control group (t = -1.40, $p = .16$). This finding is important as it demonstrates that the ratio affects perceived credibility independent of the content of the social media posts.

6.4. Discussion

Study 1 used a controlled setting and revealed that a high ratio decreases the perceived credibility of social media posts compared to posts that have a low ratio (H1) or that contain no information about the ratio at all. results also showed that our model is robust against the content type (i.e., marketing tweets vs. news).

7. Study 2

Study 1 showed that the comment-to-like ratio of a social media post affects how credible consumers judge that post. However, Study 1 did not rule out the possibility that

either the number of comments or likes alone could cause this effect (instead, as hypothesized, the ratio).

Study 2, has the objective of demonstrating that the ratio (and not the comments or likes alone) affects consumers' perceived credibility. Study 2 also aims to shed initial light on the process by which the ratio affects the perceived credibility of social media posts—which we hypothesized to be the perceived degree of consensus (H2). It is noteworthy that Study 2 uses more subtle manipulations of the ratio than Study 1 and a self-created social media platform (instead of one that resembles Twitter to generalize across platforms). The study was pre-registered at aspredicted.org (see Web Appendix 9).

7.1. Procedure

We recruited 1,002 participants (M_{Age} = 40.7, SD_{Age} = 12.6, 40.6% male) from MTurk in exchange for a monetary payment. Participants of the previous study were not eligible to participate in this study. Study 2 used a 2(ratio: high vs. low) × 5(format: indicators available to participants; see description of conditions below) between-subjects design.

First, participants learned about a new social media platform. They received information that this platform had similar characteristics that resembled current social media platforms. After familiarizing participants with the unique characteristics of the new social media, we presented them with a manipulated post from a company introducing their new smart bottle that can measure water intake throughout the day.

Participants were randomly assigned to either the 'high' or 'low' ratio condition. In the high ratio conditions, the social media post had 15 comments and 5 likes. In the low ratio condition, the same post had 5 comments and 15 likes. Thus, the sum of the number of comments and likes was equal across the conditions.

We also manipulated the format of indicators that participants could see. Participants in the 'number of comments' condition were shown only the number of comments of the

posts. Participants in the 'number of likes' condition only learned about the number of likes. In the 'number of comments and likes' condition, participants learned about both the number of comments and likes. In the 'number of comments, likes, and ratio' condition, consumers not only learned about the number of comments and likes, but also about the ratio. In the 'ratio' condition, participants learned only about the ratio, but did not receive any information about the underlying number of comments and likes (see Web Appendix 2).

As the motivation of this study was to show whether—as hypothesized—the ratio or the individual comments affected perceived credibility, we collapsed the conditions in which participants only learned about either comments or likes to 'No Ratio.' We also collapsed the conditions in which participants explicitly learned about the ratio to 'Explicit Ratio.' The remaining condition (in which participants learned about both comments and likes) was the 'Implicit ratio.' This allowed us to test whether the ratio and not just the number of comments or likes in isolation, was what mattered for perceived credibility. Web Appendix 10 shows additional analyses of the conditions.

7.2. Measurement

We measured the perceived credibility as in previous studies (Cronbach a = .96). Perceived consensus was measured using a single-item Inclusion of Other in the Self (IOS) scale (Aron et al., 1992). To measure perceived consensus, participants had to indicate the extent to which they believed the comments on the post agreed with its content on a scale from 1 to 7. The scale contains two circles that participants were told represent people and the account, and they were instructed that the closer the circles are, the more consensus there is between them regarding the post's content. We additionally assessed participants' perceptions of the scale of likes, comments, and the ratio of each item using distinct one-item Likert-type scales. Web Appendix 11 summarizes the descriptive statistics of the study.

7.3. Results

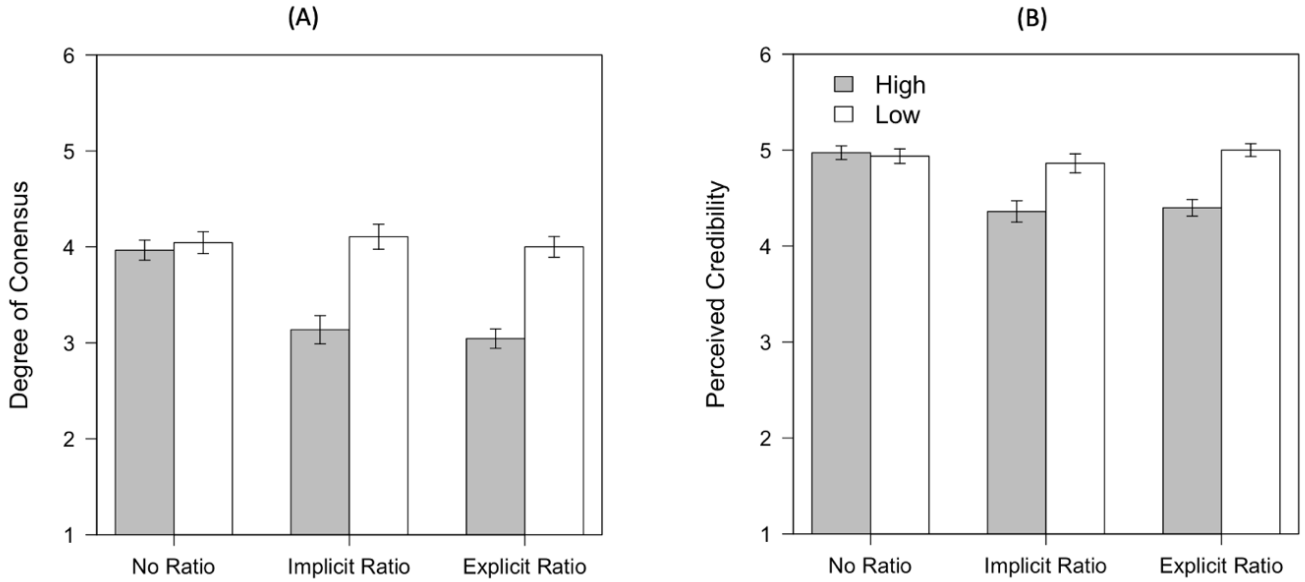
7.3.1. Degree of consensus

A two-way ANOVA with the ratio and the collapsed conditions as independent variables and degree of consensus as the dependent variable was significant ($F_{Interaction}(2, 996) = 10.43$, $p < .001$, $\eta^2 = .02$, $F_{Ratio}(1, 996) = 43.29$, $p < .001$, $\eta^2 = .04$, $F_{CollapsedFormat}(2, 996) = 10.41$, $p < .001$, $\eta^2 = .02$, $M_{ExplicitRatio\_High} = 3.05$, $SD = 1.43$, $M_{ImplicitRatio\_High} = 3.14$, $SD = 1.43$, $M_{NoRatio\_High} = 3.97$, $SD = 1.50$, $M_{ExplicitRatio\_Low} = 4.00$, $SD = 1.55$, $M_{ImplicitRatio\_Low} = 4.11$, $SD = 1.38$, $M_{NoRatio\_Low} = 4.04$, $SD = 1.54$). Planned contrasts revealed that the effect of the high ratio was significant in the explicit ($t = -6.48$, $p < .001$) and implicit ratio ($t = -4.69$, $p < .001$) conditions, but not in the no ratio conditions ($t = -.52$, $p = .61$). Thus, a high (vs. low) ratio affected the degree of consensus only in conditions that explicitly or implicitly (i.e., revealing both comments and likes) revealed the ratio, but not in conditions that revealed individual indicators (i.e., either comments *or* likes).

7.3.2. Perceived credibility

Another two-way ANOVA using the ratio and the collapsed conditions as independent variables and perceived credibility as dependent variable was significant ($F_{Interaction}(2, 996) = 9.76$, $p < .001$, $\eta^2 = .02$, $F_{Ratio}(1, 996) = 24.87$, $p < .001$, $\eta^2 = .02$, $F_{CollapsedFormat}((2, 996) = 8.36$, $p < .001$, $\eta^2 = .01$, $M_{ExplicitRatio\_High} = 4.40$, $SD = 1.22$, $M_{ImplicitRatio\_High} = 4.36$, $SD = 1.08$, $M_{NoRatio\_High} = 4.97$, $SD = 1.02$, $M_{ExplicitRatio\_Low} = 5.00$, $SD = .96$, $M_{ImplicitRatio\_Low} = 4.86$, $SD = 1.05$, $M_{NoRatio\_Low} = 4.94$, $SD = 1.03$). Follow-up contrasts revealed that the effect of the high ratio was significant in the explicit ($t = -5.72$, $p < .001$) and implicit ratio ($t = -3.40$, $p < .001$) conditions, but again not in the no ratio conditions ($t = .33$, $p = .74$). Thus, a high (vs. low)

ratio only affected perceived credibility in conditions that explicitly or implicitly revealed the ratio, but not in conditions that revealed individual indicators (Figures 3A, B).



Figs. 3. The Effect of a High (vs. Low) Ratio on (A) Degree of Consensus and (B) Perceived Credibility Depends on the Format of the Ratio.

7.3.3. Mediation analysis (no ratio)

A mediation analysis using model 4 of Preacher Hayes with ratio as independent variable, degree of consensus as mediator, perceived credibility as dependent variable revealed that for the 'no ratio' condition, there is no significant indirect (B = -.01, SE = .03, CI95% = [-.07; .04]) or direct (B = .05, SE = .10, CI95% = [-.15; .25]) effect.

7.3.4. Mediation analysis (implicit ratio)

Another mediation analysis using model 4 of Preacher Hayes with ratio as the independent variable, degree of consensus as mediator, and perceived credibility as the dependent variable showed that for the 'implicit ratio' condition, there is a significant indirect effect (B = -.21, SE = .07, CI95% = [-.36; -.10]) while the direct effect (B = -.29, SE = .15, p = .05, CI95% = [-.59; .01]) is marginally significant.

7.3.5. Mediation analysis (explicit ratio)

Another mediation analysis using model 4 of Preacher Hayes with ratio as independent variable, degree of consensus as mediator, perceived credibility as dependent variable showed that for the 'explicit ratio' condition, there is significant indirect (B = -.24, SE = .05, CI95% = [-.34; -.15]) and direct effect (B = -.36, SE = .11, CI95% = [-.57; -.15]).

7.3.6. Moderated mediation analysis

A final moderated mediation analysis using model 4 of Preacher Hayes with ratio as the independent variable, degree of consensus as mediator, perceived credibility as dependent variable, and whether the participants had access to the ratio or not as moderator, revealed a significant indirect effect for the ratio present conditions (B = -.22, SE = .04, CI95% = [-.30; -.16]), while the indirect effect for the ratio absent conditions was insignificant (B = -.02, SE = .04, CI95% = [-.09; .05]). The index of moderated mediation was also significant (B = -.21, SE = .05, CI95% = [-.31; -.11]).

7.4. Discussion

Study 2 revealed that the comment-to-like ratio, and not the individual numbers of comments or likes, drives the degree of consensus and perceptions of the credibility of a post. More precisely, participants who only learned about comments or likes (instead of both) did not report a greater or lower degree of consensus or perceived credibility while those who learned about the ratio (implicitly or explicitly) did so.

Interestingly, explicitly calculating the ratio increases the effect of a high (vs. low) ratio above and beyond not doing so. However, also the implicit ratio is sufficient to affect perceptions of credibility. This is suggestive that consumers automatically, or heuristically calculate the ratio, and no effortful calculation or evaluation takes place.

This study also identified perceived consensus as a potential cause of the variations in perceived credibility. Specifically, a high ratio decreased the degree of consensus and eventually perceived credibility of a social media post.

Finally, the design of this study (i.e., using a self-created social media platform and more conservative ratios) had several implications: First, the effects should be generalized across platforms. Second, the ratios did not rely on participants' prior comprehension of the magnitude of comments or likes on commonly used social media platforms (e.g., on Twitter).

We conducted another study (see Web Appendix 3) that rules out other potential mediation paths for the model such as ambiguity, interest in the post, positive and negative feelings toward the post, and involvement with the topic.

8. Study 3

The previous studies provided the first indications that the degree of consensus explains why the ratio affects the perceived credibility of social media posts. However, the evidence was purely correlational and did not afford causal conclusions.

Study 3 was designed with the objective of providing causal evidence on whether degree of consensus indeed causally mediated the effect of the ratio on perceived credibility. This study also provides an initial test of the downstream consequences of the ratio by testing whether the ratio and perceived credibility of social media posts ultimately also affect purchase intention for products included in those posts. This study was pre-registered at aspredicted.org (Web Appendix 9).

8.1. Procedure

We recruited 403 participants (M_{Age} = 45.9, SD_{Age} = 13.3, 44.6% male) from MTurk in exchange for a monetary payment that all participants who completed the study received.

Those who participated in one of the previous studies could not participate. This study used a 2(ratio: high vs. low) × 2(consensus cue: present vs. absent) between-subjects design.

At the beginning of the study, participants saw a timeline of a non-branded technology account. The timeline included three social media posts about different new brands of smartwatches (see Web Appendix 5). Participants learned that the social media posts represented the most recent reviews of the products that the account posted. Notably, the content of the tech account resembled actual descriptions of the products on Amazon.com and thus was realistic. We also informed participants that all three smartwatches resembled in terms of quality and price.

We manipulated the ratio of the second tweet as in the previous studies. In the 'high ratio' conditions, the ratio was almost 208 (i.e., 807 comments and 4 likes). In the 'low ratio' conditions, the ratio was 1/208 (i.e., 4 comments and 807 likes).

We manipulated the degree of consensus cue as follows. In the 'consensus absent' conditions, participants did not receive any additional information. In the 'consensus present' conditions, participants learned that most of the comments on the manipulated post supported the content of the post. Thus, the 'consensus absent' condition resembled those of prior studies, while the 'consensus present' conditions should increase the degree of consensus cue. We expect that the consensus attenuates the effect of the ratio on the perceived credibility.

8.2. Measurement

Participants rated the perceived credibility of each of the social media posts using the same scale as in the previous studies (Cronbach a = .96).

We assessed purchase intention by a one-item measure ("How likely would it be for you to purchase the stated smartwatch in the future?" ranging from 1 = "Unlikely" to 7 = "Very likely").

Moreover, perceived degree of consensus cue, age, gender, and experience level in using social media were measured as in Study 2. Web Appendix 11 shows the descriptive statistics.

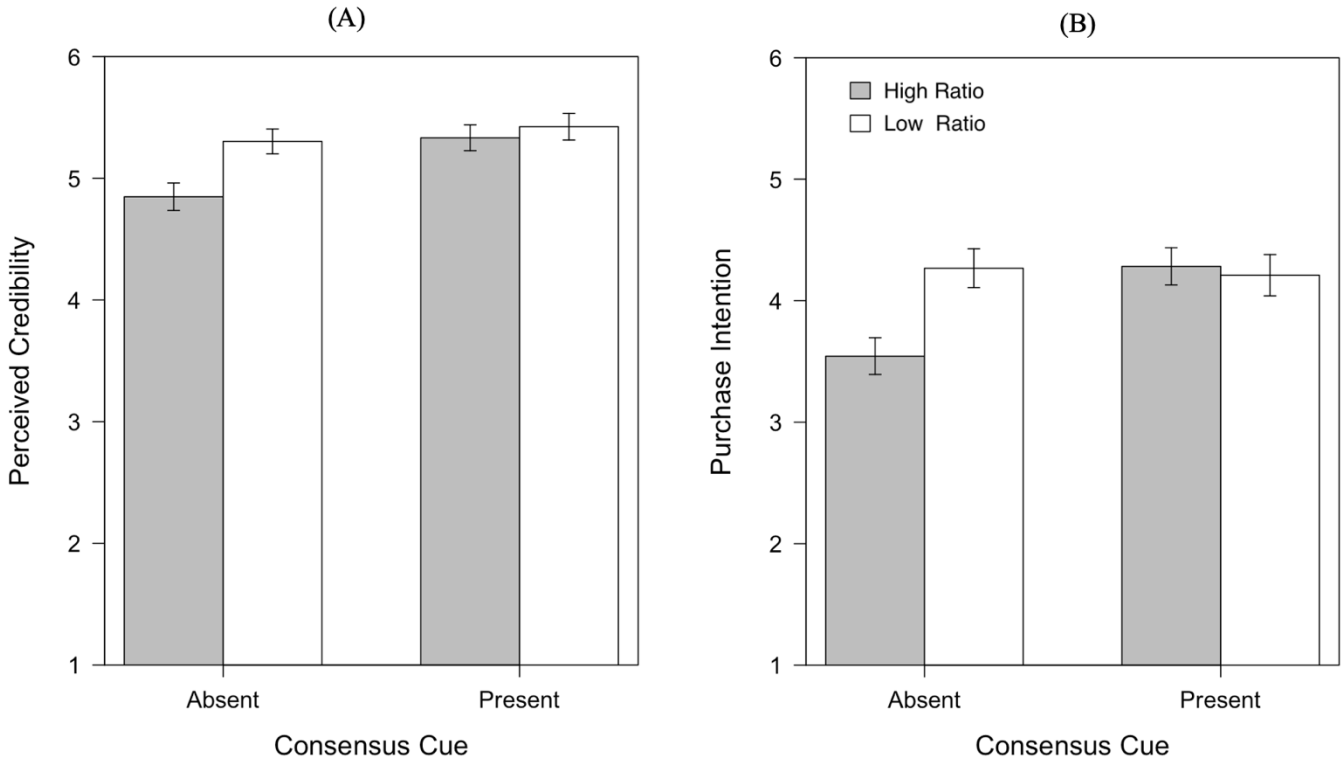
8.3. Results

8.3.1. Manipulation check (perceived degree of consensus cue)

A two-way ANOVA with consensus cue and ratio as independent variables and perceived degree of consensus cue as dependent variable ($M_{HighRatio\_AbsentConsensus}$ = 3.54, SD = 1.60, $M_{HighRatio\_PresentConsensus}$ = 4.85, SD = 1.34, $M_{LowRatio\_AbsentConsensus}$ = 4.26, SD = 1.44, $M_{LowRatio\_PresentConsensus}$ = 4.62, SD = 1.49) showed that the manipulation of consensus cue had a significant effect ($F(1, 399) = 32.45$, $p < .001$, $\eta^2 = .07$) on perceived degree of consensus cue. Thus, the manipulation worked as intended.

8.3.2. Perceived credibility

A two-way ANOVA with consensus cue and ratio as independent variables and perceived credibility as dependent variable ($M_{HighRatio\_AbsentConsensus}$ = 4.85, SD = 1.15, $M_{HighRatio\_PresentConsensus}$ = 5.33, SD = 1.06, $M_{LowRatio\_AbsentConsensus}$ = 5.30, SD = 1.03, $M_{LowRatio\_PresentConsensus}$ = 5.42, SD = 1.09) revealed a marginally significant interaction ($F(1, 399) = 2.84$, $p = .09$, $\eta^2 = .01$) and significant main effects of ratio ($F(1, 399) = 6.62$, $p = .01$, $\eta^2 = .02$) and consensus cue ($F(1, 399) = 7.90$, $p = .005$, $\eta^2 = .02$). In the absent conditions, the ratio had a significant negative effect on perceived credibility ($t = -3.00$, $p = .003$), while in the consensus present conditions this effect was non-significant ($t = -.59$, $p = .55$). Thus, a high (vs. low) ratio again reduced (increased) the perceived credibility of the focal social media post, while providing a consensus made this effect vanish. The results are shown in Figure 4.



Figs. 4. The Effect of the Ratio on (A) Perceived Credibility and (B) Purchase Intention is Moderated by Degree of Consensus Cue.

8.3.3. Purchase intention

A two-way ANOVA with consensus cue and ratio as independent variables and purchase intention as dependent variable ($M_{HighRatio\_AbsentConsensus}$ = 3.54, SD = 1.53, $M_{HighRatio\_PresentConsensus}$ = 4.28, SD = 1.53, $M_{LowRatio\_AbsentConsensus}$ = 4.27, SD = 1.61, $M_{LowRatio\_PresentConsensus}$ = 4.21, SD = 1.70) revealed a significant interaction ($F(1, 399) = 6.30$, $p = .01$, $\eta^2 = .02$). Moreover, the main effect of ratio ($F(1, 399) = 4.40$, $p = .04$, $\eta^2 = .01$) and consensus cue ($F(1, 399) = 4.64$, $p = .03$, $\eta^2 = .01$) were significant. In the consensus absent conditions, the ratio had a significant negative effect on purchase intention ($t = -3.24$, $p = .001$), while in the consensus present conditions, this effect was non-significant ($t = .32$, $p = .75$). Thus, a high (vs. low) ratio again reduced the purchase intention for the product shown in the social media post, while providing a consensus cue made this effect vanish.

8.4. Discussion

This study demonstrated that manipulation of the consensus cue can moderate the mediation of the impact between the ratio and perceived credibility, thereby confirming that the findings of Study 2 are not affected by the endogeneity of perceived consensus in our model. These results offer compelling causal evidence that the degree of consensus is responsible for the effect of the ratio on perceived credibility. Moreover, the study revealed that perceived credibility has a significant influence on purchase intention.

9. Study 4

The previous studies revealed that a high ratio decreased the perceived credibility of the social media post. However, they did not show whether this effect is directly relevant for marketers. Prior literature showed that the perceived credibility of online ads affects purchases (e.g., Kim & Song, 2020; Muda & Hamzah, 2021; Flanagin et al., 2014; Floh et al., 2009).

Study 4 was designed to test whether the ratio not only affected perceived credibility of a social media post and purchase intention, but also purchase behaviors in an incentive-compatible setting. Moreover, this study used more realistic stimuli and a more realistic experimental setting to increase the ecological validity of the results.

9.1. Procedure

We recruited 1,000 participants (M_{Age} = 40.0, SD_{Age} = 12.9, 46.3% male) from MTurk in exchange for a monetary payment that all participants who completed the study received. Participants of previous studies were excluded.

This study used a 2(ratio: high vs. low) between-subjects design. After the introduction and the attention check at the beginning of the study, participants were initially presented with a timeline of a user that featured three tweets related to three distinct new

brands of smartwatches (see Web Appendix 6). Participants learned that the account discusses technology-related news, and these three tweets represented their most recent reviews. The content of the tweets reflected the products' actual descriptions (as seen on Amazon.com). Participants also learned that the price ($200) and quality of all products were similar. We also informed participants that one of them would receive the product that they chose in the study as a bonus after completing the study—making the design incentive-compatible. This design closely mimicked the actual setting in which consumers are exposed to tweets, specifically Twitter and its timeline.

Controlling the quality and price of the brands, the ratio (number of comments and number of likes) of the second of the three tweets was manipulated to arrive at a high (vs. low) ratio. In the 'high ratio' condition, the ratio was 90.22 (number of comments = 812, number of likes = 9). In the 'low ratio' condition, the ratio was 9/812 (number of comments = 9, number of likes = 812).

9.2. Measurement

The credibility of each tweet was evaluated by participants using the same scale as in prior studies (Cronbach $\alpha = .95$). To assess purchase behavior, we measured product choice intention by asking participants to select one of the three options available in the tweets as their preferred choice. To do so in a consequential manner, we informed participants that one of them (selected at random) would receive their chosen product or the monetary equivalent as a bonus. The order of all measures was randomized. We assessed the susceptibility to social influence using a 4-item scale on a range from 1 (indicating "Totally Disagree") to 7 (indicating "Totally Agree"), as established by Stöckli & Hofer (2020) with a Cronbach's α reliability coefficient of .88. This assessment was conducted at the conclusion of the study, aimed at examining the potential impact of this construct on our subsequent investigation. Placing this assessment after the survey questions and before demographic inquiries was a

deliberate choice, designed to emulate a real-life consequential event. Age, gender, and experience level in using social media were also measured as before. Web Appendix 11 shows the descriptive statistics.

9.3. Results

9.3.1. Perceived credibility of the focal tweet

We first examined whether the perceived credibility of the focal tweet differed across conditions. A t-test with conditions as the independent variable and perceived credibility as the dependent variable showed that a high ratio significantly decreased the perceived credibility ($t(998) = -5.27$, $p < .001$, Cohens d = -.33, $M_{HighRatio} = 5.20$, SD = 1.14, $M_{LowRatio} = 5.55$, SD = .94). Thus, a high (vs. low) ratio again reduced the perceived credibility of the focal tweet.

9.3.2. Perceived credibility of all tweets

Next, we examined whether the effects were specific to the focal tweet with the manipulated ratio, or, whether the effects even spilled over to the perceived credibility of the other two tweets. A linear mixed model with conditions as the independent variable and perceived credibility as the dependent variable and the tweet (focal tweet vs. non-focal tweets) as moderator for all tweets showed that a high ratio decreased the perceived credibility of the tweet significantly ($B_{HighRatio} = -.35$, SE = .06, $t = -5.75$, $p < .001$). For the tweets whose ratios were not manipulated, the moderation was significant and positive ($B_{HighRatio \times NonFocalTweet} = .28$, SE = .04, $t = 6.89$, $p < .001$); Thus, only the credibility of the focal, manipulated tweet was affected (see Figure 5).

9.3.3. Choice intention

Only focusing on the focal tweet, a logit regression with conditions as the independent variable and choice intention as the dependent variable showed that a high ratio decreased the product choice intention of the tweet significantly (B = -.44, SE = .13, $z = -3.47$, $p < .001$).

Thus, a high (vs. low) ratio again reduced the product choice intention for the product shown in the focal tweet. Moreover, repeating the same model including the control variables such as susceptibility to social influence and level of expertise in social media, replicated the results (B = -.44, SE = .13, z = -3.44, $p < .001$).

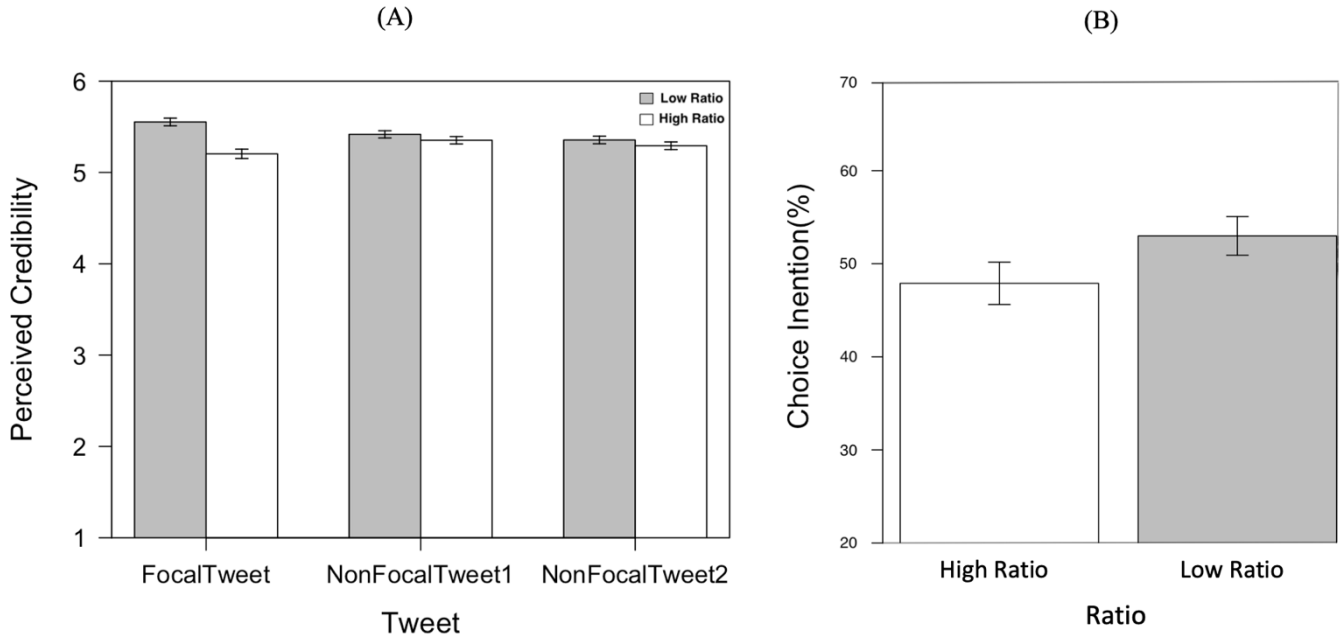


Fig. 5. (A) The Ratio of the Focal Tweet Primarily Affects the Perceived Credibility of the Focal Tweet (And Not of Other Tweets) and (B) the Percentage of the Participants Who Chose the Focal Tweet

9.3.4. Mediation by perceived credibility

Next, we analyzed whether perceived credibility mediated the effect of ratio on purchase. Using the PROCESS Model (Hayes, 2012; Model 4, $N_{Bootstraps}$ = 10,000) with ratio as independent variable, perceived credibility as mediator, and choice intention of the focal tweet as the dependent variable, revealed a significant direct effect (B = -.32, SE = .1, CI95% = [-.58; -.07]) and significant indirect effect (B = -.10, SE = .04, CI95% = [-.19; -.01]) of our manipulation on choice intention. Thus, perceived credibility partially mediated the effect of ratio on product choice intention.

9.4. Discussion

This study provides strong evidence that the ratio has a practically relevant effect. Specifically, the results document that a high ratio not only decreases perceived credibility of a social media post but also purchase intention for a product that is presented in a social media post. Notably, this study assesses consumer purchase behavior in a consequential manner and by using a design that closely resembles a realistic environment: the timeline of Twitter. Taken together, this study demonstrates the downstream consequences of the ratio of social media posts in a setting of high ecological validity.

10. Study 5

The studies presented so far showed that the comment-to-like ratio affects perceived credibility of social media posts and demonstrated its significance to marketers. However, it is unclear how to mitigate the negative impact of ratioing on perceived credibility for posts that may receive a substantial number of comments (i.e., posts regarding a controversy surrounding the brand). Therefore, the objective of Study 5 is to test whether the content type, in particular, requesting feedback in the social media post, could attenuate the negative effects of ratioing on potentially controversial content. This study was pre-registered at aspredicted.org (Web Appendix 9).

10.1. Procedure

We recruited 401 participants (M_{Age} = 40.5, SD_{Age} = 12.4, 49.1% male) from MTurk in exchange for a monetary payment that all participants who completed the study received. Participants of the previous studies were excluded.

Participants were randomly assigned to one of the four conditions of the 2 (high vs low ratio) × 2(feedback: solicited vs. not solicited) between-subjects experiment.

In the first stage, participants were introduced to a new social media platform identical to the one used in Study 2. After acquainting participants with the social media platform, they saw a social media post from a company that introduced a new smart bottle (same as in Study 2).

In the 'high ratio' conditions, the social media posts had 15 comments and 5 likes, while in the 'low ratio' conditions, the same posts had 5 comments and 15 likes. Thus, the sum of likes and comments was equal for both conditions. Participants in the 'no feedback solicited' conditions saw the normal post about the smart bottle. Participants in the 'feedback solicited' conditions saw the same social media post with the exception that a request for feedback was included at the end of the post "Please share your feedback about our smart bottle in the comment section" (see Web Appendix 7).

10.2. Measurement

Participants rated perceived credibility of the tweet (Cronbach a = .96) with the same 9-item scale used in other studies. We include a one-item manipulation check for feedback solicitation, asking participants about their agreement to the statement "The post was asking for feedback from customers" (from 1= "Totally disagree" to 5= "Totally agree"). Eventually, age, gender, and level of experience in using social media were measured as before.

10.3. Results

10.3.1. Manipulation check

A two-way ANOVA with feedback and ratio as independent variables and perceived feedback solicitation as dependent variable ($M_{HighRatio\_AbsentFeedback}$ = 2.24, SD = .99, $M_{HighRatio\_PresentFeedback}$ = 3.73, SD = 1.14, $M_{LowRatio\_AbsentFeedback}$ = 2.25, SD = .83, $M_{LowRatio\_PresentFeedback}$ = 3.72, SD = 1.26), showed that the feedback solicitation conditions resulted in greater perceptions of feedback solicitation ($F(1,397) = 194.59$, $p < .001$, $\eta^2 = .33$). The manipulation worked as intended.

10.3.2. Perceived credibility

A two-way ANOVA with feedback solicitation and ratio as independent variables and perceived credibility as dependent variable ($M_{HighRatio\_AbsentFeedback}$ = 4.35, SD = 1.24, $M_{HighRatio\_PresentFeedback}$ = 4.58, SD = 1.22, $M_{LowRatio\_AbsentFeedback}$ = 5.17, SD = .86, $M_{LowRatio\_PresentFeedback}$ = 4.99, SD = 1.07) revealed a marginally significant interaction on perceived credibility ($F(1, 397) = 3.47$, $p = .06$, $\eta^2 = .01$). Neither the ratio ($F(1, 397) = 32.11$, $p < .001$, $\eta^2 = .01$) nor soliciting feedback had a significant effect ($F(1, 399) = .06$, $p = .80$, $\eta^2 < .001$). Follow-up contrasts showed that a high ratio significantly reduced perceived credibility in the no feedback solicitation conditions (B = -.82, t = -5.42, $p < .001$), but this effect for the feedback solicitation conditions was significantly smaller (B = -.40, t = -2.44, $p = .02$). The results are shown in Figure 6.

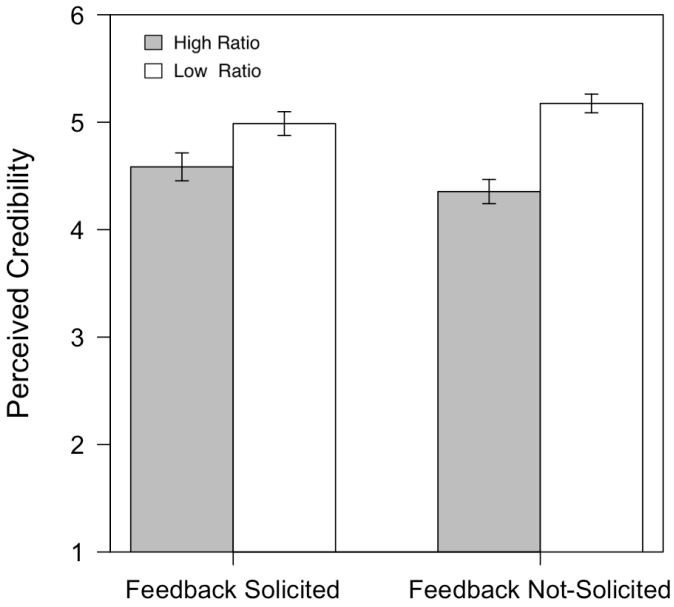


Fig. 6. The Effect of Ratioing Being Attenuated After Asking for Feedback in the Content of the Post.

10.4. Discussion

This study demonstrates that marketers can potentially mitigate the negative impact of ratioing on perceived credibility by soliciting feedback in the post. This way consumers may

interpret the number of comments—and thus the comment-to-like ratio to be less negative (i.e., less in terms of disagreement).

11. General discussion

This paper examined the phenomenon of 'ratioing' in social media, which describes the situation when a social media post receives a (significantly) greater number of comments than likes. Six studies—one field study and five online experiments—reveal that the ratio between the number of comments and likes has a significant effect on the perceived credibility of a social media post and that this perceived credibility eventually affects consumer purchase behavior. The studies also demonstrate that this effect is caused by consumers using the ratio as a cue for the perceived degree of consensus of others with the content of the social media post.

The results show that the ratio of comments to likes affects the perceived credibility of a post and not the magnitude of the individual numbers of comments or likes in isolation. Whether the calculated ratio was explicitly presented or only implicitly through the individual numbers of comments and likes did not affect the perceived credibility of a post. If a post receives 10 or 1000 comments is not sufficiently informative to judge consensus because the number of exposures is not known. The number of likes not only provides a benchmark against which the number of comments can be judged as high or low, but it also provides information about the valence of the reaction. Consequently, a post featuring 100 comments and 1000 likes should be perceived as equally credible as a post with 1000 comments and 10000 likes, all else equal.

This paper also reveals a potential strategy for mitigating the negative effects of high comment-to-like ratios by changing the content framing strategy of the social media post. In particular, by soliciting feedback for a ratioed post, consumers perceive the post to be less

negative (in terms of perceived credibility), by changing the way that users may interpret the ratio as a heuristic cue. While this strategy could not entirely prevent the negative effect of a high ratio, it reduced the effect. Future work could complement the current findings and test different other strategies to mitigate the effect of the ratio.

It is important to note that the term "ratioing" is commonly used to describe a social media post that has received more comments than likes, with a ratio of 1 being the threshold for such a post to be considered "ratioed." However, in this paper, our focus is not on this specific definition of "being ratioed." Instead, we concentrate on the ratio between the number of comments and likes as a continuous metric.

Finally, it is important to mention that even while several of our studies used manipulations that resembled Twitter, there are significant indications that the effects documents are more general and apply to other social media platforms. We created a social media platform for Studies 2, 5, and the study reported in Web Appendix 8 to show that the observed effects generalize to other platforms—as long as they reveal both the number of comments and likes. This practice is commonplace, with platforms like Facebook, LinkedIn, Instagram, TikTok, and Twitter all following similar approaches.

11.1. Theoretical contributions

This paper makes several contributions to the growing literature on how consumers process information on social media. Prior research revealed that consumers use characteristics of the source, such as the profile picture (Xu, 2014), or specific characteristics of a post like grammar and punctuation (Morris et al., 2012), as well as the ratio between the number of accounts who follow an account and the number of accounts that are followed by an account (Valsesia et al., 2020) to judge the content of a social media post. In this paper, we identify a social cue by which consumers assess the content of a post. We show that the ratio affects the degree to which consumers perceive a post to be credible from the mediation

channel of perceived degree of consensus. In essence, it functions as a peer-review process and thus accounts for the assessment of fellow consumers' judgments in an aggregated way (i.e., the number of comments and the number of likes that a post receives). The findings of Study 2 are particularly noteworthy because they demonstrate that the number of comments or likes alone do not affect the perceived credibility of a social media post; it is the comment-to-like ratio that is critical in assessing credibility.

This paper also provides strong evidence about the mechanism behind the effect of the ratio on the perceived credibility of a social media post. In particular, the results show that the perceived degree of consensus operates as a mediator of the effect of the ratio on perceived credibility. Prior work on the perceived degree of consensus on social media is sparse. Lee (2015) showed that perceived consensus is theoretically relevant but not significantly related to the number of likes alone. The results of our studies, in contrast, show that it is the ratio between comments and likes and not the absolute number of likes that drives perceived consensus.

Finally, identifying the level at which the ratio starts backfiring was not the objective of this paper and future research should fill this gap. However, the Study in Web Appendix 3 provides a first indication of this level. It finds that starting at a ratio of about eleven consumers begin to judge the credibility of a social media post as significantly lower. Nevertheless, Studies 2, 5, and the study in Web Appendix 8 show that perceived credibility is negatively and significantly affected at a much lower ratio. This contributes to how the ratio functions in the context of social media as it provides a first indication of the specific threshold while prior work on ratios lacks revealing such concrete ratio levels. While only one single study tests various levels of the ratio, this finding is consistent with our argument that consumers assume that a positive comment would most likely be accompanied by a like.

Thus, a greater number of comments than likes would imply that those comments are negative. Further research is needed to confirm this assumption.

11.2. Practical contributions

The insights of this article can directly be used by practitioners who use social media to communicate about their companies or products. First, our studies show that even a ratio that is based on a small number of comments and likes affects consumers' judgment of the credibility of a social media post. This is significant because firms can use this information to evaluate the credibility of their posts, even when they receive little consumer feedback (e.g., comments and likes). For example, a brand can monitor the ratios of each post as a general metric of its social credibility, along with other metrics such as impressions and the number of likes. This can help the company to continuously control and adjust its content marketing strategies on social media in response to the aggregated reaction of its real audiences. If a particular type of post receives more comments than likes (more than the critical threshold), for instance, the brand may decide to replace that specific type of content. Additionally, this metric can serve as a marketing research tool on a macro level, allowing companies to understand the perceived credibility of different segments of users towards new technologies, brands, or ideas on social media.

Study 4 demonstrates that the ratio and the perceived credibility eventually could promote purchases of the product mentioned in the post, which highlights the practical relevance of the revealed mechanism. Study 5 indicates that content-framing strategies can directly impact how people perceive ratioing as a cue for perceived consensus. In essence, marketers can use feedback-seeking strategies for controversial posts to reduce the strength of the ratioing cue. This strategy makes it more likely that users will perceive the number of comments as indicative of positive discussion rather than disagreement. In that way, the study demonstrates a boundary condition of ratioing on perceived credibility.

We also conducted an additional study (refer to Study Web Appendix 8) to evaluate the impact of the content of comments themselves on consensus and credibility, as opposed to solely their quantity. The study's findings demonstrate the robustness of our model against the comments shown below the post, indicating that a positive first comment (which is typically the only comment automatically displayed to users) fails to counteract the negative impact of ratioing. This emphasizes the significance of marketers and PR teams being proactive, rather than reactive, in managing potentially controversial posts and considering implementing the 'feedback seeking' strategy beforehand.

While the key hypothesis of this paper was that the comment-to-like ratio affects the perceived credibility of a social media post, we are aware of arguments for reverse causality in practice. Consumers could make comments and likes based on their judgments of the credibility of those posts. Nevertheless, it is important to distinguish between perceived credibility (which is subjective) and credibility (or factuality, which can be objective). For example, a fake advertising post may prompt individuals with knowledge of the topic (who helps them to label the content as not true) to comment in disagreement without liking the post, resulting in an initial pool of comments and likes that could influence the perception of the post's credibility among those unaware of the issue. Even if we assume that reality is subjective (and, therefore, there is no difference between perceived credibility and actual credibility), we still face a simultaneous causal relationship between the ratio and perceived credibility, which makes the phenomenon a dynamic system. Although our model does not explain how the system reaches the point of being ratioed, could be a valuable area for future research, it examines the downstream effects of such ratioing on other users, and our findings suggest that this effect is significant.

In sum, this paper examines the consequences of the number of comments and likes that a post receives. Future research could examine other responses by consumers, such as

retweets, along with their consequences on consumer judgments. It is also worthy studying the effect of ratioing on social media platforms with different 'liking' structures, for example on Amazon when we have scaled ratings (ranging from 1 to 5), or for LinkedIn when we can have different kinds of reactions, with different valence, to the post. Furthermore, researchers could examine additional boundary conditions of the revealed effect. For instance, one could examine whether the ratio has more or less impact on the perceived credibility of a social media post given a source that might already be disputable or not. Finally, it is worth noting that ratioing can have additional implications for consumers' attitudes toward the brand or advertising. For example, if ratioing undermines users' sense of autonomy or freedom in relation to a post, it may motivate them to respond to the social media post (Brehm, 1966), such as by reading the comments or researching the topic, to reconcile any perceived incongruency in their assessment of the post's credibility.

Do not limit the comments! The effect of restricting the ability to comment on social media, on suspicion and attitude toward the brand of users.

ABSTRACT

This paper examines the repercussions of undermining one of the central characteristics of social media, i.e., restricting people's ability to comment on social media posts. The authors hypothesize that social media posts that do not allow others (e.g., followers) to comment on those posts may potentially harm consumers' perceptions of the brand making that post. Precisely, consumers could consider social media posts that do not allow for commenting, as a tool for suppression of freedom of speech which eventually cultivates negative attitudes towards the brand or person that posted that specific post. A combination of a field study and seven online experiments substantiate this claim. The studies demonstrate that the impact of closing the comment section on attitude towards the brand is mediated by perceived censorship, which heightens user suspicion. We also show that this relationship is robust across various contexts, including the importance of the topic and whether the author consistently prohibits comments as a communication strategy for all the posts. This work provides novel insights into a recent phenomenon—restricting consumers' ability to comment on social media posts—and how consumers respond to this. The results also furnish actionable insights to companies for refining their communication management strategies in the digital environment.

Keywords: Social Media Indicators, Perceived Censorship, Suspicion, Information Processing on Social Media, Attitude Toward Brands.

1. Introduction

Social media are characterized by their interactive, bi-directional nature. This characteristic allows companies to engage with their current and prospective consumers while also serving as a user-friendly way for people to share their opinions and experiences with the online world (Dellarocas, 2003). Despite this openness, recent strides in social media—across platforms like Twitter, YouTube, and Instagram—have introduced a new dimension: the option for users to restrict the ability to comment on their posts.

While the reason for granting this option from social media platforms is to allow for greater control over privacy, it is unclear whether and how consumers respond to their fellow users or companies using this option. Thus, the question "What are the consequences of such actions on users' attitudes?" arises.

Consumers constantly process information from a wide array of channels (Jang et al., 2017), whether for receiving news, following their favorite brands, or even making purchase decisions. What is unique to social media platforms and thus sets them apart from conventional media is its multifaceted role for companies, serving not solely as a communication platform (e.g., about their brand and products), but also to engage with their consumers. This paradigm shift has redefined the interactions between brands and users, affording both parties increased insight into one another, thereby altering the landscape (Sashi, 2012).

Moreover, social media have empowered consumers to directly voice themselves through actions such as crafting content, engaging with existing posts, or utilizing the comments and views of other users under a post as a reference for potential brand purchases, thus reconfiguring the very essence of their relationship (Westerman et al., 2014). Consequently, consumers are not just users of these platforms, but also content generators for each other. User-generated comments support developing more genuine characteristics about

a brand's offerings, reducing potential information asymmetry between consumers and companies, and thereby increasing the transparency of the consumer-brand association (Xiao et al., 2018). Furthermore, the ability to comment allows consumers to ask questions or raise concerns, which allows brands to excel in responsiveness and interactivity, both of which can increase the perceived credibility of the responses of the brand (Le et al., 2021).

In other words, the ability to comment on social media posts serves as an important cue that allows consumers to express their opinions but also facilitates positive judgments about the person or brand that allows for those comments. Users can make decisions that are closer aligned with their preferences and needs. Furthermore, the ability to comment on social media posts also has a crucial role also for forming purchase decisions (Algi, 2018). The interactive nature of social media creates a sense of community engagement and offers the ability to exchange insights, and opinions, and to receive advice from their peers, which can eventually lead to customer loyalty and satisfaction (Brodie et al., 2013).

As the platforms evolve, recent developments such as the ability to limit or restrict comments on posts introduce a layer of control on privacy, yet also prompt an inquiry into the potential consequences of such choices. This paper explores whether restricting users' ability to comment on a social media post could have adverse consequences. Specifically, it investigates the potential repercussions of using this option on post-engagement and consumers' overall brand perception. By investigating these dynamics, we aim to shed light on the evolving interplay between brands, consumers, and the ever-shifting landscape of social media communication.

2. The psychology of restricting the ability to comment on social media posts

One of the key functions of social media is the ability to interactively exchange, and therefore also to respond (i.e., "comment") to one another's posts. While restricting the ability

to comment on social media posts might have both positive (e.g., enhanced privacy, avoidance of negative content) and negative consequences, we propose that consumers' intuitive responses are adverse—they tend to perceive any restrictions to comment on others' posts as an act against the freedom of speech. In a nutshell, for those who use the commenting function, that function is valuable, independent of the purpose (e.g., engaging with brands, gaining insights), as it serves as a medium for users to exchange their ideas and opinions on various topics.

The concept of freedom of speech is closely intertwined with individualism and independence, standing in contrast to the specter of potential government intervention (Riedl et al., 2021). Given the prevailing norm on social media platforms that users can comment publicly, we propose that users could consider any restrictions as a form of censorship, potentially translating into a perceived loss of their freedom of expression (Oh & Aukerman, 2013), thus making restricting comments as an even more general signal of censorship. The significance of scrutinizing this phenomenon intensifies, as numerous individuals regard freedom of speech as their paramount right in society (Shugan, 2006).

Censorship is basically defined as stopping someone from expressing their thoughts or speech, which were presumed to be free or not restricted before (Freshwater, 2004). Censorship encompasses a wide array of actions, not just the deliberate concealment or manipulation of information, which often involves disrupting communication channels, making them inaccessible to users, as highlighted by Varol and Uluturk (2019). Censorship can be harmful for many reasons, as it can allow powerful organizations to control what information gets out, hiding truths that might be unpopular which can hurt free markets by limiting choices and competition (Shugan, 2006).

Thus, according to the definition of censorship, restricting the ability of users (or fellow consumers) to comment on a post can be considered an act of censorship, just as

deleting a piece of news or someone's review can be. This is because it disrupts communication channels and probable information that could have been provided in the comment section. However, the question remains whether users perceive this act as censorship, or more as a protective measure for privacy and information.

To address this question, a preliminary, exploratory investigation ($N_{MTurkers}$ = 100, M_{Age} = 36.3, SD_{Age} = 10.8, 59% male) took place which exposed participants to a social media post of a brand introducing a product. Notably, the ability to comment was visibly restricted. Participants then shared their perceptions of the brand in an open text box. Text analysis revealed that this practice made them suspicious about the content of the post, felt a manipulative intent of the brand (e.g., removal of negative comments), and expressed their perceptions of censorship (e.g., impeding users from expressing their views), even though some users stated that the reason for this act could be related to preempting hate speech or, in extreme cases, trolling behavior from social media users, potentially safeguarding oneself, the brand, or other users, from reputational harm. Appendix 18 provides additional details on this study.

In the context of our Pre-Study, it becomes clear that most users primarily interpret the restriction of comments as a signal for brands to assert control over information. This cue emerges even in the absence of direct evidence suggesting such control, supporting the idea that censorship is not only removing or hiding information, but also the interruption of the communication channel, which here is the ability to comment. This unintended consequence of censorship can obstruct open dialogue and, notably, foster suspicion among users.

Utilizing cues and signals is not a new concept in decision-making. Recent research showed that individuals extensively rely on informational cues and signals in their decision-making processes on social media. These cues encompass various elements that indirectly convey the credibility of information within the social media realm. For instance, individuals

assess a user's authenticity based on factors like their profile picture (Xu, 2014), the presence of a verified badge, and even specific aspects of their posts, such as grammar and punctuation (Morris et al., 2012). Moreover, individuals also consider nuanced social indicators, such as a user's follower-to-following ratio (Valsesia et al., 2020) or the ratio of comments to likes on a post (Rezaee et al., TBD), when forming judgments about a user or their posts in the social media sphere.

Considering that individuals adeptly leverage these nuanced cues for their decision-making processes, it is plausible to deduce that users may similarly interpret a brand's decision to disable comments on a post as a significant cue that influences their judgments, regardless of their intention to comment on the post. The lack of comments, once symbolic of open engagement, has the potential to morph into a trigger that leads users to view the content as edited or questionable. Consequently, this paper aims to thoroughly examine this phenomenon, delving into whether restricting post comments can indeed alter user attitudes by manipulating perceptions of censorship and suspicion.

Suspicion, a product of uncertainty and triggered by perceived deceptive actions, plays a crucial role in shaping customer perceptions and responses. As Burgoon et al. (1996) elucidate, suspicion revolves around a belief held without concrete evidence, hinting at potential dishonesty in an individual's speech or actions. It embodies a lingering doubt that questions the authenticity of presented information, often arising from the identification of deceptive traits within an action. Researchers like Bart et al. (2005) and Zhuang et al. (2018) emphasize how perceived deceptive attributes can sow doubt around an action, prompting individuals to approach it with heightened skepticism.

Deception, especially in the realm of communication, occurs when an individual intentionally attempts to instill false beliefs, as explained by DePaulo and DePaulo (1989). This entails manipulating information to establish or perpetuate beliefs that the communicator

knows to be untrue. Similarly, the act of censoring information or disrupting conversations can exert a similar influence on user behavior and may be considered a form of deception (Varol & Uluturk, 2019). Such manipulation extends to the domain of marketing, where dishonest actions or messages can influence individuals' purchasing decisions by encouraging belief in unverifiable claims or by sowing uncertainty about the buying process (Aditya, 2001).

Consumer suspicion, as defined by Buller and Burgoon (1996), revolves around the uncertainty regarding the truthfulness of a sender's message. This suspicion, fueled by enduring uncertainty arising from perceived deception, creates an environment in which the credibility of presented information is questioned. Oza et al. (2010) underscore the role of consumer awareness in instigating suspicion and promoting vigilance against deceptive tactics, especially within product reviews.

Moreover, the proliferation of technology and social media has concurrently witnessed an escalation in deceptive practices by brands. The mounting suspicion among users toward online environments can be attributed to several key factors, leading to the perception that restricting comments may be more deceptive than protective. Notably, recent observations have unveiled the employment of varied strategies by sellers to mitigate the impact of negative reviews, as elucidated by Hendy (2019). These strategies include manipulating the visibility and sentiment of reviews and comments, such as concealing unfavorable reviews, and potentially distorting perceptions of product quality. As users become more attuned to these maneuvers, their inherent skepticism toward the authenticity of online content deepens. Additionally, media coverage and practitioner studies have underscored a disconcerting trend where unknown digital brands accumulate unverified reviews on platforms like Amazon, as highlighted by Woollacott (2019). This surge of potentially unauthenticated reviews not only

dilutes the credibility of online testimonials but also fuels reservations concerning the genuineness and trustworthiness of digital information.

Zhuang et al. (2018) shed light on the predisposition of individuals who have fallen victim to online fraud to question the accuracy and credibility of information presented in product reviews. These prior experiences with deceitful practices contribute to a prevailing wariness that taints their perception of online content, extending to measures like restricting the comments. Thus, interpreting the restriction of the comments becomes increasingly crucial, as more individuals may encounter censored information, inevitably evaluating subsequent limited comments with heightened skepticism.

Thus, another salient factor that amplifies the significance of scrutinizing the restriction of comments is its incongruence with the typical norms of social media posts. In essence, when brands opt to post content with restricted comments, they deliberately deviate from the prevalent pattern. This deliberate divergence signals that there is an underlying motive behind this unconventional approach. Regrettably, the prevailing climate of skepticism towards social media posts and digital information at large complicates users' ability to readily attribute this action to a positive and authentic rationale. Instead, the prevailing climate of skepticism inclines users to lean towards interpreting this measure as an act of censorship, driven by motives other than safeguarding the online environment.

In amalgamation, these references underscore the evolving landscape of digital mistrust, where deceptive practices, the influx of unverified information, personal encounters with online fraud, and manipulative seller conduct coalesce to render users markedly skeptical of online environments. Within this context, users are inclined to perceive the restriction of the comments as an act of deception rather than protection, aligning with their heightened apprehension towards the authenticity of digital content.

H1: Restricting consumers' ability to comment on social media posts (vs. not) increases consumers' perceived suspicion about the post.

H2: The impact of restricting consumers' ability to comment on social media posts on suspicion is mediated through the perceived degree of censorship.

3. Reactance theory and attitude toward the brand

The perception of limited comments as indicative of censorship holds the potential to trigger intricate dynamics in users' attitudes. Anchored in congruent theory, when individuals confront incongruent attitudes, they are inclined to rectify this incongruence. One plausible way of achieving this congruency is by seeking additional information (Osgood and Tannenbaum, 1955). Consequently, the restriction of people's ability to comment on a social media post may prompt consumers' suspicion. This, in turn, could potentially result in reduced purchase intentions and a less favorable view of the brand (DeCarlo, 2005). Hence, if there is indeed something to conceal, it might be more beneficial to make it transparent to users rather than keeping it hidden.

Expanding the framework, reactance theory (Brehm, 1966) furnishes a lens to comprehend user responses to restricted comments. If users construe the closure of comments as an endeavor to suppress dissenting opinions, they may be propelled to resist such censorship and reclaim the liberty of expression curtailed by this action. In this context, users could counteract the comments' restriction by generating adverse word-of-mouth, drafting negative reviews, engaging in unfavorable quote retweeting, and other analogous behaviors. The application of reactance theory adds a layer of depth to the understanding of user conduct in the wake of restricting the ability to comment, particularly within the realm of social media platforms. This theoretical construct serves as an invaluable tool to explore the intricate dynamics underlying user reactions to the curtailed interactive space. It affords the potential

to elucidate how individuals navigate the intersection of their perceived freedom of expression and the platform's regulatory mechanisms, yielding insights into the broader ramifications of such actions on user engagement and discourse.

Furthermore, the discontent stemming from perceived freedom loss, or perceived censorship, might extend beyond the platform with restricted comments. Reactance theory suggests that individuals may actively seek alternative avenues to express their opinions and regain a sense of control (Brehm, 1966). In such instances, users who have been restricted to one platform might be inclined to share their negative experiences on uncensored platforms. This could materialize as the dissemination of negative word-of-mouth on these alternative platforms. This reactance effect can sometimes significantly undermine the intended objective of censorship. For instance, as noted by Rappaport (1997), attempts to censor a book or movie can paradoxically increase users' inclination to read or watch them.

H3: The impact of restricting the ability to comment on users' suspicion can decrease the user's attitude toward the brand.

However, the interplay of diverse contextual factors, encompassing variables like the subject matter's consistency, the historical context, and the perceived importance of the topic, could potentially moderate these effects (Eckstein, 1997), as they can explain the motive behind such acts. Therefore, it is plausible that the consistency of the act of restricting the comments (whether applied to all of a brand's posts or not), the level of importance attributed to the topic (whether there might be a significant reason for censoring the information), and the historical context of the brand (whether the brand has faced issues like trolling, hate speech, or misleading comments in the past) may change the impact of restricting comments on users' behavior and perception. These factors contribute to the complexity of user reactions to the absence of comment sections, highlighting the intricate web of variables that shape users' perceptions and responses in the digital landscape.

4. Overview of studies

We conducted a total of eight studies—one field study and seven pre-registered online experiments—to examine our hypotheses and assess the impact of restricting the ability to comment on users' brand attitudes. The field study provides initial evidence of the relationship between restricting the ability to comment and users' attitudes using social media posts (N = 45,435 tweets). Study 1 (N = 202) tested the core hypothesis through a controlled experiment, revealing that restricting the comments diminishes brand attitude by amplifying perceived censorship and suspicion. Study 2 (N = 202) meticulously explored the mediation channel of perceived censorship, while effectively ruling out other potential mediators like consistency, customer responsiveness, and trolling history. This study solidified that the analyzed effect indeed stems from the mediation channel of perceived censorship. Study 3 (N = 400) manipulated the level of censorship to furnish *causal* evidence for its role as the primary mediator between closing the comments and suspicion/brand attitude. Study 4 (N = 499), Study 5 (N = 401), and Study WA1 (N = 400) delved into theoretically proposed moderators that could mitigate or counteract the negative effect of closing the comments on users' attitudes. These three studies underscored the model's robustness against parameters such as "level of importance of the topic," "within-account consistency of the brand's action in closing comments on social media," and "providing a legitimate historical context to explain the rationale behind closing the comments." Finally, Study 6 (N = 501) demonstrates that not only is restricting the ability to comment after receiving negative replies worse than doing so before posting, but it also reveals that having negative replies is preferable to restricting comments after receiving such feedback. You can find the link to the pre-registered studies in Appendix 13.

5. Field study

The purpose of this field study is to conduct an initial assessment of the correlation between the closure of comments on social media posts and consumers' attitudes toward the post. This investigation focuses on the primary hypothesis of this research within a naturalistic setting, ensuring a high level of ecological validity.

5.1. Procedure

In this study, a total of 49,981 English-language social media posts were randomly collected from Twitter during August 2023. The primary objective was to uncover any significant correlation between the tweet status and the attitude of the response that the tweet received from other users. Specifically, we used the quote-retweets that the tweets received (whether they conveyed negativity or positivity in the quotes) to assess the attitude of the users toward the tweets.

Since the posts with restricted comments didn't have any direct comments attached to them, we had to rely on quote-retweets of the tweets as the user responses to those tweets. To do this, we gathered all the quote-retweets from tweets with closed comments (totaling N = 328). We then utilized the assistance of two RA to evaluate the attitude toward the main tweet of these quotes. To provide a basis for comparison, we also collected N = 232 random quote-retweets from normal (control) tweets from the same database. This approach allowed us to assess how users' responses in quote-retweets of posts with closed comments differed from those of typical tweets, shedding light on the impact of comment restrictions on user engagement and attitude.

5.2. Measurement

5.2.1. Attitude

The attitude of the quote-retweets toward the original tweet was assessed with the help of two research assistants, blind to the hypothesis, who coded the last 20 quote-retweets of each tweet separately. In our study, research assistants were tasked with evaluating the attitude of quotes in relation to the original tweets. Using a 0-10 scale, assistants rated quotes based on their disposition towards the content of the original tweet, rather than the general tone of the comment itself. A score of 0 indicated complete disagreement with the original tweet, 10 represented full agreement, and 5 signified neutrality. It was emphasized that alignment with the tweet's content could result in a high score, even if the content was negative (e.g., agreeing with a tweet that dislikes a movie, with a negative tone, would result in a positive attitude). In cases where the language of a quote was not understood, a neutral score of 5 was assigned.

The coding from the two research assistants showed a high correlation ($r = .82$, $p < .001$), with an intraclass correlation coefficient of .81 ($F(388,389) = 9.25$, $p < .001$), underscoring the validity of the labeling process. The final measurement for assessing user attitudes toward the tweets was derived from the average of the scores provided by the two assistants.

5.2.2. Content Type

We also used the help of a research assistant to categorize the tweet content aiming to control for any confounding effects related to the topic of each tweet. Tweets were labeled as *ads* (promotional content, whether commercial or personal), *opinions* (comments on various topics or other tweets), *news* (discussions directly related to news events, whether verified or not), or *daily-tweets* (general commentary on the individual's day-to-day life, without a specific focus).

Appendix 14 provides a summary of the descriptive statistics for the field study and the subsequent experiments.

5.3. Result

In our analysis of a subset of 389 tweets with coded quote-retweets, we conducted a linear regression using the average attitude score of quote-retweets as the dependent variable and the tweet's status (Restricted vs. Control) as the independent variable. Our findings indicate that restricting the ability to comment on a tweet significantly lowers the attitude expressed towards it in quote-retweets (B = -.64, SE = .18, t = -3.58, $p < .001$). Even when controlling for variables such as content type, verification status, and the account's follower-to-following ratio, the results remained consistent (B = -.77, SE = .18, t = -4.20, $p < .001$). These findings were also replicated when using the coding from either the first or the second research assistant individually (first RA: B = -.49, SE = .21, t = -2.36, $p = .02$; second RA: B = -1.05, SE = .18, t = -5.93, $p < .001$). Thus, the evidence strongly suggests that restricting comments harms the sentiment expressed in quote-retweets toward the original tweet.

Furthermore, our argument, grounded in reactance theory, posits that restricting the ability to comment has the potential to amplify negative word-of-mouth interactions among users, creating a sense of congruency. To investigate this further, we harnessed the complete dataset, comprising 49,981 collected tweets. We employed an alternative regression model with the natural logarithm of the number of quotes as the dependent variable, given the substantial skewness in the distribution of this variable (skewness = 47.9, $p < .001$). The independent variable in this model was the tweet's commenting status. In addition, we controlled for various factors, including content type, verification status, and the account's follower-to-following ratio. The results of this regression analysis revealed a significant increase in the number of quote retweets when the ability to comment was restricted (B = .26, SE = .05, t = 4.80, $p < .001$). This finding provides initial evidence supporting the notion that limiting comments can indeed foster greater negative word-of-mouth interactions among users.

5.4. Discussion

This field study offered an initial insight into the negative correlation between restricting the ability to comment on social media posts and consumers' attitudes toward the post. Specifically, the data show that the tweets with restricted comment sections receive more negative quotes than the tweets with not restricted comments. As with any field study, this investigation is not without limitations. The study utilized uncontrolled field data within a correlational design, which inherently precludes the establishment of causality. Our assessment of attitudes toward brands relies on correlational metrics, which can potentially introduce measurement errors specific to this construct. Nonetheless, we employed this study to underscore the correlated associations among our focal parameters using real-world data. In upcoming studies, our emphasis will shift towards investigating causality more comprehensively.

6. Study 1

The primary goal of Study 1 is to explore the potential impact of comment restrictions on attitudes towards the brand. This will be achieved by investigating whether such limitations lead to an increase in perceived censorship and suspicion among users. To achieve this objective, an experimental design will be employed to manipulate the comment section of the posts under examination.

6.1. Procedure

We recruited 202 respondents (M_{Age} = 40.0, SD_{Age} = 12.7, 47.0% male) from Amazon Mechanical Turk ("MTurk") in exchange for monetary compensation. This study was conducted using a between-subjects design featuring two distinct conditions (Restricted vs. Control). At the study's commencement and following the introduction, we gauged participants' value priorities employing a modified version of Rokeach's original Value Survey (1968). Specifically, participants were tasked with ranking the provided values: (1) "A Comfortable Life," (2) "A Sense of Accomplishment," (3) "A World at Peace," (4) "Equality," (5) "Family Security," (6) "Freedom," (7) "Inner Harmony," and (8) "Wisdom." This measurement was implemented to ascertain whether the emphasis placed on freedom within participants' value systems could influence perceived censorship and its subsequent impact on suspicion levels.

Subsequently, participants were presented with a tweet originating from a fictional brand named "Kheyzaran," promoting their latest smartwatch to their target audience (Appendix 15). These participants were randomly divided into two groups: one exposed to the normal version of the tweet (control version with open comment section), and the other exposed to the tweet with a restricted comment. Following their exposure to the tweet, participants were prompted to evaluate their perceived censorship of the post, levels of

suspicion, and attitudes towards the brand. Participants viewed the assessment in a randomized order.

6.2. Measures

6.2.1. Attitude toward the brand

After the exposure to the manipulations, participants were tasked with evaluating their attitude towards the brand. To gauge this, we employed a modified version of the 5-item scale of Attitude toward the Brand in the Ad, adapted from the work of Lee et al. (1999). (Cronbach a = .85).

6.2.2. Suspicion

We assessed participants' levels of suspicion using an adaptation of a 9-item scale originating from the work of Obermiller et al. (1998), specifically designed to measure skepticism towards advertising (Cronbach a = .93).

6.2.3. Perceived censorship

Perceived censorship was evaluated using a straightforward 1-item Likert-type scale. Participants were prompted to indicate their level of agreement with the statement, "The brand (Kheyzaran) was trying to censor information from users," utilizing a scale ranging from 1 (Totally disagree) to 7 (Totally agree).

6.2.4. Manipulation check

We assessed the efficacy of our manipulation by directly inquiring about the number of comments the observed posts received from a selection of four available options: 10, 230, 75, and the indication that the post's comments were restricted.

6.2.4. Demographics

Preceding the demographic inquiries, participants engaged in addressing two hypothesis guess questions. In an open-ended format, participants had the opportunity to speculate about the underlying hypotheses. Additionally, they could select from six potential

answers regarding the study's purpose. Notably, all provided options were plausible, such as "the influence of post grammar on consumers" or "the effect of the number of likes on consumers." Concluding the survey, all participants furnished information about their age (grouped), gender, and their familiarity with using social media. You can find the descriptive statistics of this study in Appendix 14.

6.3. Results

6.3.1. Manipulation check

We conducted a one-way ANOVA using our manipulation (Restricted vs. Control) as the independent variable and the manipulation check as the dependent variable ($M_{Restricted}$ = 3.79, SD = .68, $M_{Control}$ = 2.75, SD = 1.33). The results indicate that our manipulation was effective in the study ($F(1, 200) = 49.7$, $p < .001$, $\eta^2 = .2$).

6.3.2. Direct effects

6.3.2.1. Perceived censorship

Utilizing a one-way ANOVA ($F(1, 200) = 36.40$, $p < .001$, $\eta^2 = .15$), with our manipulation serving as the independent variable and perceived censorship as the dependent variable, we observed a significant effect in relation to the restriction of the comments. The mean perceived censorship was higher in the 'Restricted' condition (M = 3.89, SD = 1.70) compared to the 'Control' condition (M = 2.58, SD = 1.37), indicating that closing the comments indeed led to an increase in perceived censorship by the participants.

Furthermore, our interest extended to examining whether the incorporation of freedom as a value priority and the users' social media experience could potentially alter the impact of our model. To address this, we employed a linear regression approach. Here, the manipulation served as the independent variable, freedom priority was introduced as the moderator, and the degree of experience with social media was introduced as a control variable. The results of this analysis demonstrated the persistence of the primary effect (B = .94, SE = .45, t = 2.10, *p*

= .04), indicating that the manipulation retained its significance. However, it is worth noting that no significant moderating effect was observed for the freedom value priority (t = .93, p = .35), suggesting that the influence of freedom priority did not substantially change the relationship within the model.

6.3.2.2. Suspicion

Conducting another one-way ANOVA (F(1, 200) = 16.90, $p < .001$, $\eta 2 = .08$), utilizing our manipulation as the independent variable and suspicion as the dependent variable, revealed a significant effect of conditions on suspicion levels. Specifically, in the "Limited" condition (M = 4.28, SD = 1.06) compared to the "Control" condition (M = 3.64, SD = 1.15), closing the comment section led to a significant increase in participants' levels of suspicion.

6.3.2.3. Attitude toward the brand

Employing an additional one-way ANOVA (F(1, 200) = 7.54, p = .01, $\eta 2 = .04$), where the manipulation served as the independent variable and attitude toward the brand was the dependent variable, a significant direct influence of conditions on the dependent variable emerged. Specifically, within the "Limited" condition (M = 2.62, SD = 1.15) in contrast to the "Control" condition (M = 3.03, SD = .97), the act of closing the comment section yielded a significant reduction in participants' attitude toward the brand (Figure 7).

6.3.3. Mediation Analysis

Utilizing the PROCESS Model (Hayes, 2012; Model 6, $N_{Bootstraps}$ = 10,000), with conditions as the independent variable, attitude toward the brand as the dependent variable, and perceived censorship and suspicion as mediators, reveals a significant indirect effect between closing the comments and attitude toward the brand. This effect operates through the serial mediation of perceived censorship and suspicion (B = -.24, SE = .13, CI95% = [-.38; -.14]). Meanwhile, the indirect effects for alternative pathways, such as from manipulations to

perceived censorship to attitude toward the brand (B = -.07, SE = .06, CI95% = [-.21; .04]), and the pathway from manipulations to suspicion to attitude toward the brand (B = -.12, SE = .09, CI95% = [-.30; .05]), along with the direct effect (B = .03, SE = .13, CI95% = [-.23; .28]), remain insignificant. Moreover, changing the order of the mediators in the serial mediation, makes the mediation path insignificant (B = -.02, SE = .02, CI95% = [-.07; .01]), which again supports our hypothesis. These findings suggest the presence of a meaningful mediated pathway through perceived censorship and suspicion, influencing attitude toward the brand.

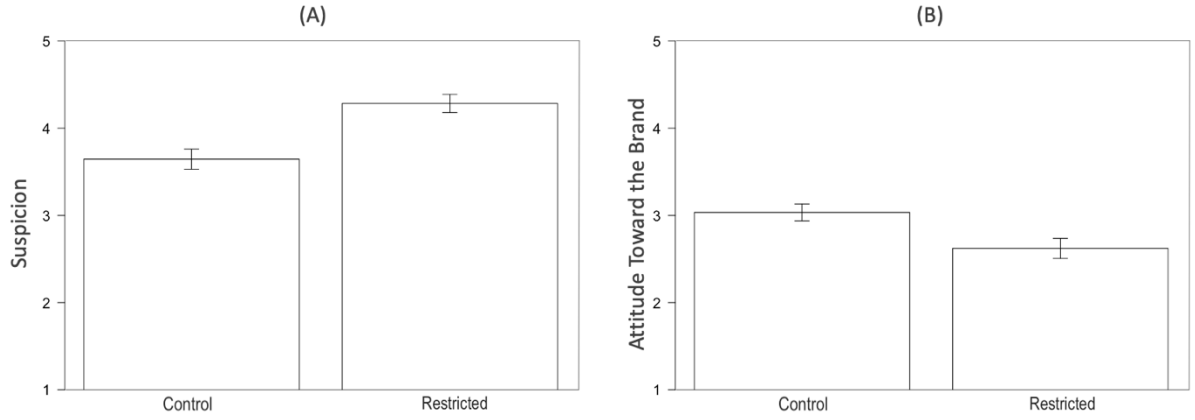


Fig. 7. The Effect of restricting the ability to comment on the post on (A) suspicion and (B) attitude toward the brand.

6.4. Discussion

The outcomes of Study 1 furnish a more heightened level of controlled and causal evidence concerning the association between the restriction of the comments and the brand attitude. Notably, this investigation substantiates the suggested sequential mediation involving perceived censorship, suspicion, and attitude toward the brand. An additional noteworthy observation from this study is that the emphasis on freedom as a value in life does not exert a moderating influence on the relationship between closing the comments and perceived censorship, which enhances the robustness of the model.

7. Study 2

Study 2's principal objective is to scrutinize the alternative potential explanations for the mediation pathway between the restriction of the comments and perceived credibility. However, it is worth mentioning that there could be also other probable mechanisms to explain restricting consumers' ability to comment on social media posts that we need to rule out. As stated before, the default on social media is that consumers are able to comment on others' posts and responses (Ahluwalia & Burnkrant, 2004). Consumer suspicion tends to rise when actions deviate from established expectations or norms. In such cases, consumers seek cues to gauge the significance of these deviations (Ahluwalia & Burnkrant, 2004) and cues to identify potential suspicious acts (Friestad & Wright, 1994; Plotkina et al., 2020). Therefore, if a brand consistently restricts consumers' ability to comment on a social media post it may be considered as the consistent communication strategy of the brand (Šerić, 2020), which may eventually outweigh the negative consequences.

Another explanation could be about the relationship dynamics between the users and the brand, and not the censoring itself. The act of closing the comments could reduce the extent to which consumers perceive a brand to be customer-responsive, thereby fostering a more negative attitude toward the brand. Perceived customer responsiveness and centricity, recognized for its influence on customer relationship performance and customer satisfaction (Habel et al., 2020; Agnihotri et al., 2016), entails a strong emphasis on addressing customers' interests (Bolton 2004; Shah et al. 2006). By curbing open dialogue and inhibiting customer participation through closed comments, brands might inadvertently convey a reduced focus on addressing customers' interests and a lesser commitment to actively engaging with their concerns which may negatively influence users' attitudes toward the brand.

The act of restricting the comments can also be interpreted as a measure to address cyberbullying, trolls, and hate speech, providing an alternative perspective to users'

perceptions. Cyberbullying, as outlined by Johnson (2011), encompasses hostile or aggressive actions carried out through information and communication technologies with the intention of harming or causing discomfort to others. In this context, by restricting the comments, brands could be seen as taking proactive steps to mitigate potential cyberbullying instances and protect users from exposure to offensive content. This perception suggests that the limitation might be a strategic approach to maintaining a safer and more respectful online environment, thus presenting a justifiable reason for users to interpret the action as an effort to curb negativity rather than an act of censorship**.**

In this study, we will examine the influence of these alternative explanations on our process model. This investigation aims to confirm that the observed effect indeed stems from the hypothesized relationship between perceived censorship and suspicion, rather than potential factors like inconsistency or a decline in perceived customer responsiveness affecting the outcomes.

7.1. Procedure

We recruited 202 participants from Amazon Mechanical Turk ("MTurk") in exchange for monetary compensation. This study employs a between-subject design (Restricted vs. Control) to investigate the potential mediating factors between the closure of comments and attitudes toward a brand (M_{Age} = 39.8, SD_{Age} = 13.0, 42.6% male). The design closely resembles Study 1, with the exception that a different fictional brand, "Aqua-Z," is utilized to promote its new fat burner (see Appendix 16). This specific manipulation was chosen to assess the impact on posts that inherently raise more suspicion than typical content.

The study commenced by having participants assess their prioritized life values, mirroring the approach taken in Study 1. Subsequently, participants were exposed to the manipulations, after which they engaged in a structured examination of multiple behavioral concepts, in a randomized order. These concepts were meticulously designed to uncover the

probable mediating mechanisms that establish a connection between the restriction of comments and the attitude toward the brand. The central aim was to methodically exclude alternative potential mediators, including customer responsiveness, consistency of action, and trolling which were introduced in the theory part. These factors were considered as they could plausibly contribute to the observed effect within the model.

7.2. Measures

7.2.1. Attitude toward the brand

In this study, we once again employed a modified version of the 5-item Attitude toward the Brand in the Ad scale, which was adapted from the research conducted by Lee et al. (1999). (Cronbach a = .92).

7.2.2. Suspicion

We assessed participants' levels of suspicion using the modified version of the 9-item scale of skepticism towards advertising of Obermiller et al. (1998) (Cronbach a = .95).

7.2.3. Perceived censorship

Perceived censorship was measured using a simple 1-item Likert-type scale as implemented in Study 1.

7.2.4. Perceived cyberbullying

We measured the perceived extent of cyberbullying targeting the user "Aqua-Z" through a 5-item scale. This scale was an adapted variant of the "perceived cyberbullying severity scale" originally formulated by Camacho et al. (2014). (Cronbach a = .97). Our primary aim was to investigate whether the act of observing a post with restricted comments could potentially suggest that the associated account had previously encountered unfavorable comments from cyberbullies, leading to the restriction of the comments. This exploration sought to discern whether the apparent reason for the closure was indicative of a response to negative interactions rather than being driven by censorship.

7.2.5. Customer responsiveness

We also measured customer responsiveness using a modified version of the 6-item scale of perceived customer centricity (Habel et al., 2020). (Cronbach a = .95). The motivation behind measuring this concept was to ensure that any adverse impact on brand attitude resulting from the restricted comments was not influenced by participants perceiving the brand as less customer-responsive. By examining this factor, we aimed to verify that the observed negative effects were not merely attributed to a reduced perception of customer-centricity in relation to the brand.

7.2.5. Consistency

Employing a 1-item Likert-type scale, we inquired whether the behavior of the "Aqua-Z" user aligned with that of other users on social media. This approach was intended to ascertain whether any inconsistencies in actions could potentially serve as an alternative explanation for the observed effect.

Conclusively, the manipulation check, along with demographic variables such as age, gender, and level of social media experience, were assessed using the same way as was implemented in Study 1. The descriptive statistics for the study can be found in Appendix 14.

7.3. Results

7.3.1. Manipulation check

We performed a one-way ANOVA, utilizing our manipulation (Restricted vs. Control) as the independent variable and the manipulation check as the dependent variable ($M_{Restricted}$ = 3.78, SD = .77, $M_{Control}$ = 2.51, SD = 1.38). The outcomes of the analysis confirm the effectiveness of our manipulation within the study ($F(1, 200) = 66.10$, $p < .001$, $\eta^2 = .2$).

7.3.2. Main mediation path

Applying the PROCESS Model (Hayes, 2012; Model 6, $N_{Bootstraps}$ = 10,000) and utilizing manipulations as the independent variable, attitude toward the brand as the

dependent variable, and perceived censorship along with suspicion as serial mediators, unveils a significant indirect effect linking the closure of the comments to attitudes toward the brand. This effect is established through the sequential mediation of perceived censorship and suspicion (B = -.57, SE = .12, CI95% = [-.81; -.36]), while the direct effect became insignificant (B = -.12, SE = .09, CI95% = [-.30; .06]).

7.3.3. Alternative explanations

By employing the PROCESS Model (Hayes, 2012; Model 4, $N_{Bootstraps}$ = 10,000) and using the manipulations as the independent variable, suspicion as the dependent variable, and perceived censorship along with customer-responsiveness, consistency, and perceived cyberbullying as parallel mediators, the results demonstrate a significant indirect effect that links the closure of the comments to suspicion. This effect is mediated by perceived censorship (B = .44, SE = .13, CI95% = [.19; .71]). Importantly, the mediation paths for consistency (B = -.003, SE = .07, CI95% = [-.15; .14]), perceived cyberbullying (B = .02, SE = .03, CI95% = [-.02; .09]), and customer-responsiveness (B = .16, SE = .09, CI95% = [-.001; .36]) remain insignificant. This outcome provides empirical support for excluding these potential explanations from our mediation model.

7.4. Discussion

Study 2 not only replicated the findings of Study 1 within a distinct context but also effectively eliminated the theoretical explanations that could have been posited to account for the observed relationship between the restriction of comments and brand attitude.

8. Study 3

The preliminary studies initially hinted at the role of perceived consensus in mediating the impact of comment restriction on suspicion and brand attitude. Nonetheless, it is important to acknowledge the potential presence of endogeneity due to the probable measurement errors

in quantifying perceived censorship. To address this concern, Study 3 was meticulously crafted with the primary aim of offering causal evidence regarding the mediating role of perceived censorship in the relationship between comment restriction and suspicion.

8.1. Procedure

We recruited 400 participants (M_{Age} = 41.6, SD_{Age} = 12.4, 47.7% male) from Amazon Mechanical Turk ("MTurk") in exchange for monetary compensation. Those who participated in one of the previous studies could not participate. Employing a 2×2 between-subjects design, this study encompassed two independent variables: Restricted vs. Control and the presence or absence of a censorship cue. In essence, our objective is to examine whether the manipulation of perceived censorship, achieved through the provision of cues indicating its presence, can moderate the impact of comment restriction on suspicion.

In this study, we adopted the identical design employed in Study 2, complete with the same manipulations. However, a novel element was introduced for participants preceding their observation of the manipulated tweet. The participants in the present cue condition were exposed to a text that conveyed the following message: "*Certain brands intentionally disable their comment sections as a means to suppress any critical or unfavorable information related to their advertisements; By doing so, they aim to maintain a carefully crafted image and shield themselves from potential backlash or public scrutiny.*" In contrast, participants in the absent cue condition were presented with the following text: "*In response to the growing concern surrounding online hate speech, brands are taking measures to create safer digital spaces for their customers; One such measure includes the decision to disable comment sections on platforms like Twitter.*"

8.2. Measures

All measurements employed in this study were consistent with those used in Study 2. These encompassed attitude toward the brand (Cronbach's α = .92), suspicion (Cronbach's α = .95), perceived censorship, age, gender, level of experience in using social media, and a manipulation check specifically related to the limitation of the comments. In addition to these, a supplementary manipulation check for the censorship cue was introduced in this study. Participants were presented with a query: "*If you come across a post with a disabled comment section from this brand, what conclusions would you draw regarding the motive behind it?*" Two response options were provided: "censoring information" and "creating a safe and private space". The descriptive statistics for this study can be found in Appendix 14.

8.3. Results

8.3.1. Manipulation checks

A two-way ANOVA was conducted, utilizing our manipulation (Restricted vs. Control) and the presence of the censorship cue as independent variables, with the censorship manipulation check as the dependent variable ($M_{Restricted\_PresentCue}$ = 1.28, SD = .40, $M_{Control\_PresentCue}$ = 1.24, SD = .43, $M_{Restricted\_AbsentCue}$ = 1.29, SD = .45, $M_{Control\_AbsentCue}$ = 1.34, SD = .48). The outcomes underscore the effectiveness of our manipulation within the study, as evidenced by the significance of the effect of the censorship cue on the manipulation check ($F(1, 396) = 4.44$, $p = .04$, $\eta^2 = .01$).

To assess the effectiveness of the censorship cue manipulation, we employed a Chi-Square test to examine its association with the corresponding manipulation check. The analysis yielded a significant result ($\chi^2 = 20$, $p = .002$), indicating that the manipulation was successful in achieving its intended effect.

8.3.2. Censorship cue

A two-way ANOVA with suspicion as dependent variable, and status of the tweet (Restricted vs. Control), and censorship cue presence as independent variable shows a significant moderating effect of censorship cue ($F(1, 396) = 6.41$, $p = .01$, $\eta^2 = .02$). Planned contrasts revealed that the effect of the comment restriction was significant in the present censorship cue ($t = -2.69$, $p = .01$), but not in the absent censorship cue condition ($t = .91$, $p = .37$).

Furthermore, utilizing the PROCESS Model (Hayes, 2012; Model 7, $N_{Bootstraps} = 10{,}000$) and incorporating the tweet condition as the independent variable, attitude toward the brand as the dependent variable, suspicion as the mediator, and censorship cue as the moderator, an interesting outcome emerges. The analysis reveals a significant index of moderated mediation for the consensus cue (B = -.52, SE = .21, CI95% = [-.93; -.12]), where the indirect effect for consensus cue absent condition was insignificant (B = .13, SE = .15, CI95% = [-.16; .43]), while the indirect effect for consensus cue present condition remains significant (B = -.39, SE = .14, CI95% = [-.66; -.11]). Thereby reinforcing the findings from the previous regression model.

8.4. Discussion

This study effectively showcased that the manipulation of perceived censorship using censorship cues possesses the capacity to moderate the influence of comment restriction on suspicion. This confirmation reinforces the robustness of the Study 2 findings, demonstrating that they remain unaffected by potential endogeneity concerns linked to perceived censorship in our model. These outcomes provide robust and compelling causal evidence, affirming that perceived censorship undeniably underpins the impact of comment limitation on both suspicion and attitudes toward the brand.

9. Study 4

This study centers on investigating the potential moderating impact of within-account consistency in restricting comments on perceived censorship and suspicion. In Study 3, we established that overall consistency—whether closing comments aligns with the behavior of other users on social media—does not exert any effect within our focal model. However, the current study focuses on a distinct phenomenon: exploring whether within-consistency, which refers to whether the user consistently employs comment restriction as part of their communication strategy or only for specific posts, influences users' perceptions of this action.

Furthermore, this study diverges by operating within a unique context—utilizing a self-designed social media platform instead of Twitter. This shift aims to scrutinize the generalizability of our hypothesis across diverse social media platforms.

9.1. Procedure

We recruited 499 participants (M_{Age} = 41.9, SD_{Age} = 13.6, 50.7% male) from Amazon Mechanical Turk ("MTurk") in exchange for monetary compensation. Individuals who had taken part in any of the prior studies were ineligible to participate. In this study, a 2×2 between-subjects design was employed, featuring two independent variables: Status (Restricted vs. Control) and the within-consistency of account in limiting comments (High vs. Low).

Initially, participants were introduced to a new social media platform and received information that highlighted its resemblances to existing platforms. Upon familiarizing participants with the distinct attributes of this new social media, including metrics like the number of likes, number of comments, and the presence or absence of restriction of comments, we proceeded to present them with a manipulated post from a company known as "XuGu GadgetWorld," which introduced their latest smartwatch.

Participants were randomly assigned to either the 'Restricted' or 'Control' comments condition. Subsequently, after viewing the post, participants were exposed to the timeline of the user "XuGu GadgetWorld," encompassing their preceding posts. Here too, participants were randomly allocated to either the 'High' or 'Low' within-consistency group. In the high consistency group, complete within-consistency was maintained across all posts' statuses, while this coherence was not sustained for the low consistency group (Appendix 17)

9.2. Measures

All measurements incorporated in this study remained in alignment with those adopted in previous studies. This encompassed attitude toward the brand (Cronbach's $\alpha = .91$), suspicion (Cronbach's $\alpha = .94$), perceived censorship, age, gender, level of experience in using social media, and a manipulation check tailored to the limitation of the comments. Furthermore, an additional manipulation check was introduced to assess the within-consistency of comment limitation. In this regard, participants were requested to evaluate their agreement with the statement: "*XuGu, the user, maintains a consistent strategy regarding their comment section on social media,"* utilizing a scale ranging from totally disagree (0) to totally agree (7). The descriptive statistics for this study can be found in Appendix 14.

9.3. Results

9.3.1. Manipulation checks

A two-way ANOVA was executed, incorporating our status manipulation (Restricted vs. Control) and within-consistency (High vs. Low) as the independent variables, while utilizing the consistency manipulation check as the dependent variable ($M_{Restricted\_HighConsistency} = 5.81$, $SD = 1.18$, $M_{Control\_HighConsistency} = 5.49$, $SD = 1.06$, $M_{Restricted\_LowConsistency} = 3.85$, $SD = 1.84$, $M_{Control\_LowConsistency} = 3.85$, $SD = 1.69$). The results showed that the consistency manipulation worked as intended ($F(1, 495) = 180.27$, $p < .001$, $\eta^2 = .27$).

Another two-way ANOVA was executed, using our status manipulation (Restricted vs. Control) and within-consistency (High vs. Low) as the independent variables, while utilizing the status manipulation check as the dependent variable ($M_{Restricted\_HighConsistency}$ = 3.60, SD = 1.00, $M_{Control\_HighConsistency}$= 1.32, SD = .85, $M_{Restricted\_LowConsistency}$ = 3.35, SD = 1.19, $M_{Control\_LowConsistency}$= 1.62, SD = 1.14). The results showed that the status manipulation also worked as intended ($F(1, 495) = 440.76$, $p < .001$, $\eta^2 = .47$).

9.3.2. Perceived suspicion

A two-way ANOVA with our status manipulation and within-consistency as the independent variables, and suspicion as dependent variable ($M_{Restricted\_HighConsistency}$ = 5.07, SD = 1.23, $M_{Control\_HighConsistency}$= 4.30, SD = 1.28, $M_{Restricted\_LowConsistency}$ = 5.15, SD = 1.04, $M_{Control\_LowConsistency}$= 4.31, SD = 1.35), shows no significant moderation effect ($F(1, 495) = .11$, $p = .74$, $\eta^2 < .001$), and direct effect of the within-consistency ($F(1, 495) = .16$, $p = .68$, $\eta^2 < .001$) on suspicion, while the main effect of post's status remains significant ($F(1, 495) = 54.05$, $p < .001$, $\eta^2 = .1$).

9.3.3. Brand attitude

Moreover, another two-way ANOVA with our status manipulation and within-consistency as the independent variables, and attitude toward the brand as dependent variable ($M_{Restricted\_HighConsistency}$ = 2.20, SD = 1.36, $M_{Control\_HighConsistency}$= 2.87, SD = 1.07, $M_{Restricted\_LowConsistency}$ = 2.09, SD = 1.15, $M_{Control\_LowConsistency}$= 2.87, SD = 1.12), shows no significant moderation effect ($F(1, 495) = 0.27$, $p = .60$, $\eta^2 < .001$), and direct effect of within-consistency effect ($F(1, 495) = 0.22$, $p = .64$, $\eta^2 < .001$) on attitude toward the brand, while the main effect of post's status remains significant ($F(1, 495) = 47.35$, $p < .001$, $\eta^2 = .09$). Thus, the results show that the model is robust against the within-account consistency of the brand in limiting the comments (Figure 8).

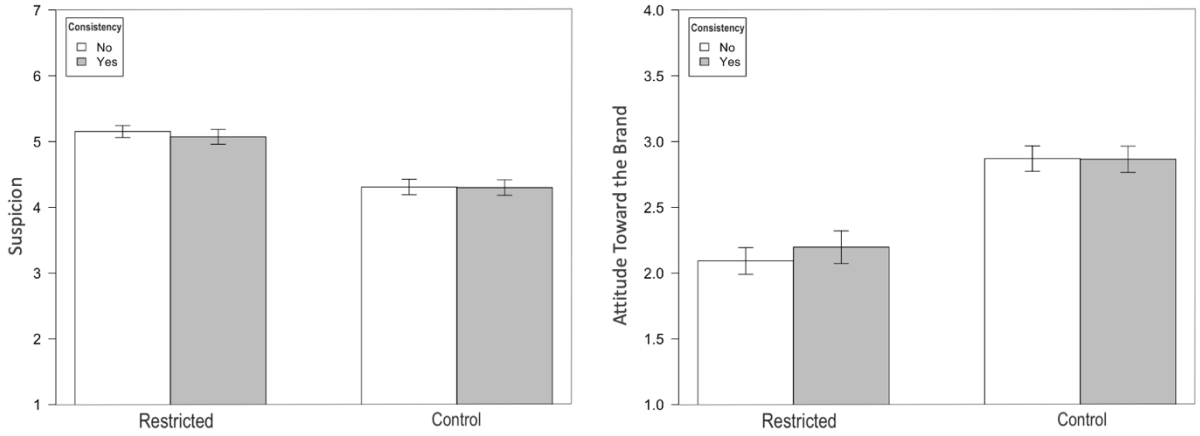


Fig. 8. The Effect of within-consistency of restricting the ability to comment of the posts on (A) suspicion, and (B) attitude toward the brand.

9.4. Discussion

This study convincingly demonstrates that the relationship between comment limitation and brand attitude remains entirely robust, regardless of the within-account consistency of brands in implementing this practice. Furthermore, the study's design, which involves the use of a self-created social media platform and employs less suspicious content for the posts, enhances the external validity and generalizability of our model, showcasing its relevance across various social media platforms.

10. Study 5

This study had the objective to examine whether explanations could moderate the negative effect of restricting the ability to comment on perceived censorship and suspicion. Specifically, this study tests whether offering a "legitimate" justification—e.g., revealing attacks and hate speech from cyberbullies—could alleviate the adverse impact of limiting the comments on consumers' brand attitudes.

10.1. Procedure

We recruited 401 participants (M_{Age} = 40.6, SD_{Age} = 12.5, 44.1% male) from Amazon Mechanical Turk ("MTurk") in exchange for monetary compensation. Furthermore, individuals who had participated in any of the previous studies were excluded from eligibility for this study. Employing a 2×2 between-subjects design, the study comprised two independent variables: Status (Restricted vs. Control) and the provision of a legitimate explanation regarding the act of limiting the comments (Absent vs. Present).

The study's design and the manipulations of the social media post and its comment section mirror those of Study 2, involving the promotion of the "Aqua-Z" brand's new fat burner. However, following the observation of the tweet, participants were randomly assigned to either the 'Present' or 'Absent' explanation group. In the present explanation group, participants were exposed to a text stating: "*We have noticed that certain brands like Aqua-Z are currently grappling with cyberbullying and hate speech on their social media. As a result, they are seeking strategies to better manage the information on their social media platforms."* Conversely, the other group did not receive any form of explanation.

10.2. Measures

All measurements implemented in this study remain consistent with those utilized in prior studies. This encompassed attitude toward the brand (Cronbach's α = .92), suspicion (Cronbach's α = .95), perceived censorship, age, gender, level of experience in using social media, and a manipulation check related to the limitation of the comments. Moreover, an additional manipulation check was introduced to gauge the efficacy of the manipulation concerning the provision of an explanation. Participants were prompted to assess their agreement with the statement: "Aqua-Z, the user, has recently experienced a distressing incident of cyberbullying," using a scale that ranged from totally disagree (0) to totally agree (7). Comprehensive descriptive statistics for this study are available in Appendix 14.

10.3. Results

10.3.1. Manipulation check

A two-way ANOVA was conducted, employing the status manipulation (Restricted vs. Control) and explanation legitimacy (Absent vs. Present) as the independent variables, and the explanation manipulation check served as the dependent variable ($M_{Restricted\_PresentExp}$ = 5.24, SD = 1.49, $M_{Control\_PresentExp}$ = 4.76, SD = 1.49, $M_{Restricted\_AbsentExp}$ = 3.02, SD = 1.25, $M_{Control\_AbsentExp}$ = 2.55, SD = 1.40). The results underscore the effectiveness of our manipulation within this study, evident through the significant influence of providing the explanation on its associated manipulation check ($F(1, 397) = 247.10$, $p < .001$, $\eta^2 = .38$), thus the manipulation of explanation legitimacy worked as intended.

Another two-way ANOVA was executed, using our status manipulation (Restricted vs. Control) and explanation legitimacy (Absent vs. Present) as the independent variables, while utilizing the status manipulation check as the dependent variable ($M_{Restricted\_PresentExp}$ = 3.81, SD = .70, $M_{Control\_PresentExp}$ = 2.16, SD = 1.40, $M_{Restricted\_AbsentExp}$ = 3.73, SD = .84, $M_{Control\_AbsentExp}$ = 2.31, SD = 1.45). The results showed that the status manipulation also worked as intended ($F(1, 397) = 181.29$, $p < .001$, $\eta^2 = .31$).

10.3.2. Moderation of explanation legitimacy

A two-way ANOVA with the status manipulation and explanation legitimacy as the independent variables, and perceived censorship as dependent variable ($M_{Restricted\_PresentExp}$ = 4.55, SD = 1.83, $M_{Control\_PresentExp}$ = 3.24, SD = 1.58, $M_{Restricted\_AbsentExp}$ = 4.98, SD = 1.74, $M_{Control\_AbsentExp}$ = 3.14, SD = 1.67), shows no significant moderation effect ($F(1, 397) = 2.24$, $p = .12$, $\eta^2 = .01$) and direct effect of explanation legitimacy ($F(1, 397) = .99$, $p = .32$, $\eta^2 = .002$) on perceived censorship, while the main effect of post's status remains significant ($F(1, 495) = 84.94$, $p < .001$, $\eta^2 = .18$). In the absent conditions, restricting the comments had a

significant effect on perceived censorship (t = 7.62, $p < .001$), while in the present explanation conditions, this effect remains significant as well (t = 5.44, $p < .001$).

A two-way ANOVA with our status manipulation and explanation legitimacy as the independent variables, and suspicion as dependent variable ($M_{Restricted\_PresentExp}$ = 4.64, SD = 1.30, $M_{Control\_PresentExp}$ = 4.54, SD = 1.27, $M_{Restricted\_AbsentExp}$ = 5.03, SD = 1.26, $M_{Control\_AbsentExp}$ = 4.70, SD = 1.43), shows no significant moderation effect of explanation legitimacy ($F(1, 397) = .77$, $p = .38$, $\eta^2 = .002$) on suspicion, while the main effect of post's status becomes insignificant ($F(1, 495) = 2.59$, $p = .11$, $\eta^2 = .01$). In the absent conditions, limiting the comments had a marginally significant effect on suspicion (t = 1.79, $p = .07$), while in the present explanation conditions this effect was non-significant (t = .55, $p = .58$).

Performing another two-way ANOVA with our status manipulation and explanation legitimacy as the independent variables, and attitude toward the brand as the dependent variable ($M_{Restricted\_PresentExp}$ = 1.96, SD = 1.37, $M_{Control\_PresentExp}$ = 2.24, SD = 1.27, $M_{Restricted\_AbsentExp}$ = 1.48, SD = 1.26, $M_{Control\_AbsentExp}$ = 1.85, SD = 1.45), unveils that there is no significant moderation effect of providing the legitimate explanation on brand attitude ($F(1, 397) = .13$, $p = .72$, $\eta^2 < .001$). However, the main effect of the post's status remains significant ($F(1, 495) = 5.63$, $p = .02$, $\eta^2 = .01$). Within the 'Absent' explanation conditions, limiting the comments significantly influenced attitude toward the brand (t = -1.98, $p = .048$), while in the 'Present' explanation conditions, this effect was non-significant (t = -1.47, $p = .14$).

10.4. Discussion

This study demonstrates that the effect of restricting the ability to respond to social media posts on brand attitude is robust, independent of a legitimate explanation for doing so. Marketers can learn from this study as it suggests that they should refrain from restricting comments, given that consumers might even ignore valid justifications. Moreover, the same

design in another study (Study WA1) shows that even the level of importance of the topic cannot have a moderating effect on our model (Appendix 19).

11. Study 6

The primary objective of this study was to investigate whether the timing of restricting the ability to comment had any impact on users' perceptions of censorship, suspicion, or their attitude toward the brand. In other words, we aimed to determine if it makes a difference whether comment restrictions are imposed before or after receiving negative comments, and, critically, whether having negative comments is more advantageous than having a post with a restricted comment section. This research holds significant implications for marketers dealing with negative feedback on their posts.

11.1. Procedure

We recruited 501 participants (M_{Age} = 39.9, SD_{Age} = 13.4, 41.7% male) from Amazon Mechanical Turk ("MTurk") in exchange for monetary compensation. Furthermore, individuals who had participated in any of the previous studies were excluded from eligibility for this study. Employing a between-subjects design, the study encompassed five distinct conditions: "Restricted before posting," "Restricted after posting without any comments," "Restricted after posting with negative comments," "Control without any comments," and "Control with negative comments."

At the outset, participants were introduced to a novel social media platform and provided with information that emphasized its similarities to well-established platforms. After acquainting the participants with the unique features of this new social media platform, such as metrics like the number of likes, number of comments, and the option to restrict comments (even after posting), we then presented them with a manipulated post from a company known as "SmartX Company," which unveiled their latest smartwatch.

Participants were randomly assigned to one of the five aforementioned conditions as follows: In the 'Restricted before posting' condition, they viewed a post in which the ability to comment was restricted prior to posting. In the 'Restricted after posting, without any comment' condition, they observed the post in which commenting was restricted after posting, and prior to any comments being made. In the 'Restricted after posting with negative comments' condition, they were presented with the post where comments were restricted only after receiving a notably negative comment. In the 'Control without any comments' condition, participants saw the post in its normal state without any comments. In the 'Control with negative comments' condition, they viewed the post that had received negative comments, yet no comment restrictions were applied (Appendix 20 for the stimuli).

11.2. Measures

All measurements utilized in this study remain consistent with those in prior studies. This included attitude toward the brand (Cronbach's $\alpha = .88$), suspicion (Cronbach's $\alpha = .93$), perceived censorship, age, gender, level of experience in using social media. Furthermore, a manipulation check was implemented to assess the effectiveness of the manipulation related to the timing of comment section restrictions. Participants were asked to select the status of the comment section of the post from among the available provided options: "The ability to comment was restricted before posting," "The ability to comment was restricted after posting," and "The comment section was open."

11.3. Results

11.3.1. Manipulation check

We conducted a one-way ANOVA using our conditions as the independent variable and the manipulation check as the dependent variable ($M_{Restricted\text{-}AfterPosting\text{-}NegativeComment} = 1.90$, $SD = .30$, $M_{Restricted\text{-}AfterPosting\text{-}NoComment} = 2.35$, $SD = .64$, $M_{Restricted\text{-}BeforePosting} = 2.89$, $SD = .40$,

$M_{Control\text{-}NegativeComment}$ = 1.29, SD = .46, $M_{Control\text{-}NoComment}$ = 1.51, SD = .82). The results indicate that our manipulation was effective in the study ($F(4, 496) = 136$, $p < .001$, $\eta^2 = .5$).

11.3.2. Perceived censorship

We conducted another one-way ANOVA using our conditions as the independent variable and the perceived censorship as the dependent variable ($M_{Restricted\text{-}AfterPosting\text{-}NegativeComment}$ = 5.36, SD = 1.45, $M_{Restricted\text{-}AfterPosting\text{-}NoComment}$ = 4.38, SD = 1.69, $M_{Restricted\text{-}BeforePosting}$ = 4.98, SD = 1.67, $M_{Control\text{-}NegativeComment}$ = 3.36, SD = 1.68, $M_{Control\text{-}NoComment}$ = 2.95, SD = 1.72). The results (Figure 9 for more details) indicate that our manipulation had a significant effect on perceived censorship ($F(4, 496) = 39.6$, $p < .001$, $\eta^2 = .24$). More specifically, the planned contrast shows that there is significant difference in perceived censorship between the condition of 'Restricted after posting with negative comments' with conditions such as 'Restricted after posting without any comments' ($t = 4.21$, $p < .001$), 'Control with negative comments' ($t = 8.59$, $p < .001$), and 'Control without any comments' ($t = 10.4$, $p < .001$). The only insignificant planned contrasts are between 'Restricted after posting without any comments' and 'Restricted before posting,' ($t = 1.63$, $p = .1$), and between 'Control without any comments' and 'Control with negative comments' ($t = 1.78$, $p = .08$), which again shows that censorship is about disturbing the communication channel, and not erasing a piece of information.

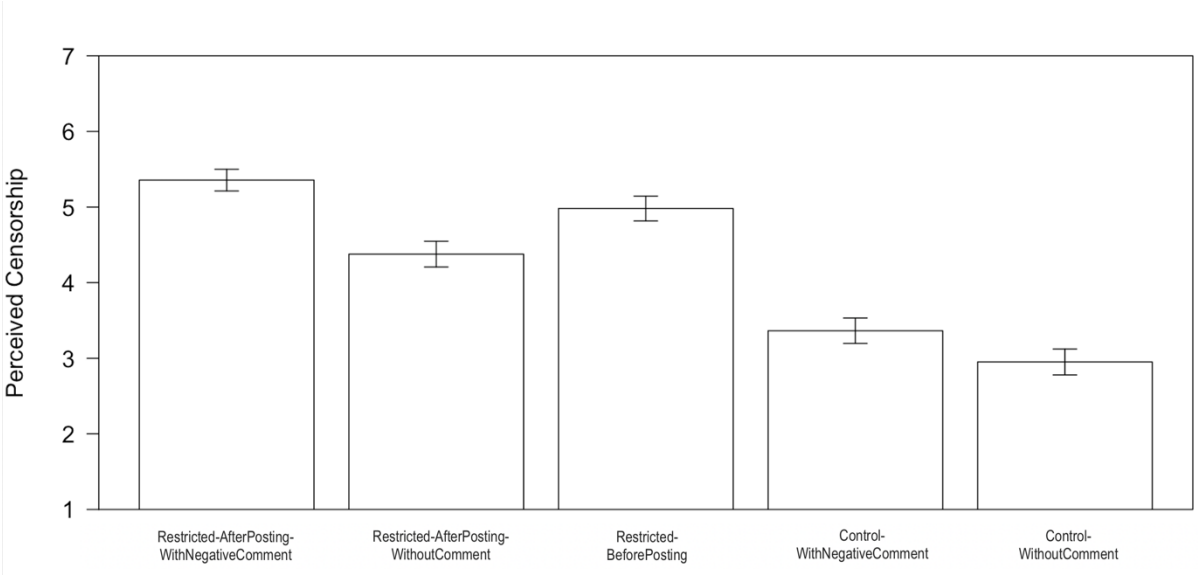


Fig. 9. The Effect of the Different Conditions on Perceived Censorship

11.3.3. Suspicion

We conducted another one-way ANOVA utilizing our conditions as the independent variable and the suspicion as the dependent variable ($M_{Restricted\text{-}AfterPosting\text{-}NegativeComment}$ = 5.02, SD = 1.05, $M_{Restricted\text{-}AfterPosting\text{-}NoComment}$ = 4.26, SD = 1.16, $M_{Restricted\text{-}BeforePosting}$ = 4.48, SD = 1.31, $M_{Control\text{-}NegativeComment}$ = 5.23, SD = .91, $M_{Control\text{-}NoComment}$ = 3.92, SD = 1.25). The results (Figure 10 for more details) indicate that our manipulation had a significant effect on suspicion ($F(4, 496) = 12.6$, $p < .001$, $\eta^2 = .1$). In greater detail, the planned contrast shows that there is significant difference in suspicion between the condition of 'Restricted after posting with negative comments' with conditions such as 'Restricted after posting without any comments' ($t = 4.67$, $p < .001$), 'Restricted before posting' ($t = 3.36$, $p < .001$), 'Control with negative comments' ($t = 3.05$, $p = .002$), and 'Control without any comments' ($t = 6.87$, $p < .001$).

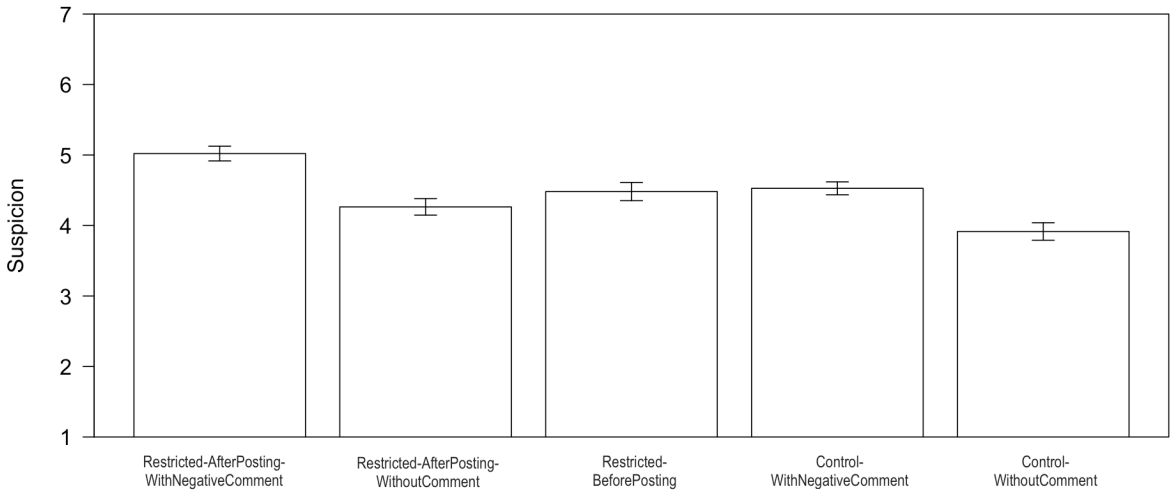


Fig. 10. The Effect of the Different Conditions on Suspicion

11.3.4. Attitude toward the brand

We utilized another one-way ANOVA utilizing our conditions as the independent variable and the attitude toward the brand as the dependent variable ($M_{Restricted\text{-}AfterPosting\text{-}NegativeComment}$ = 1.81, SD = 1.18, $M_{Restricted\text{-}AfterPosting\text{-}NoComment}$ = 2.56, SD = 1.01, $M_{Restricted\text{-}BeforePosting}$ = 2.25, SD = 1.19, $M_{Control\text{-}NegativeComment}$ = 2.09, SD = .90, $M_{Control\text{-}NoComment}$ = 2.74, SD = 1.04). The results (Figure 11 for more details) indicate that our manipulation had a significant effect on attitude toward the brand ($F(4, 496) = 12.1$, $p < .001$, $\eta^2 = .09$). especially, the planned contrast shows that there is significant difference in the measured attitude toward the brand between the condition of 'Restricted after posting with negative comments' with conditions such as 'Restricted after posting without any comments' ($t = -4.96$, $p < .001$), 'Restricted before posting' ($t = -2.95$, $p = .003$), and 'Control without any comments' ($t = -6.21$, $p < .001$), while this effect was partially significant between 'Restricted after posting with negative comments' and 'Control with negative comments' ($t = -1.88$, $p = .06$).

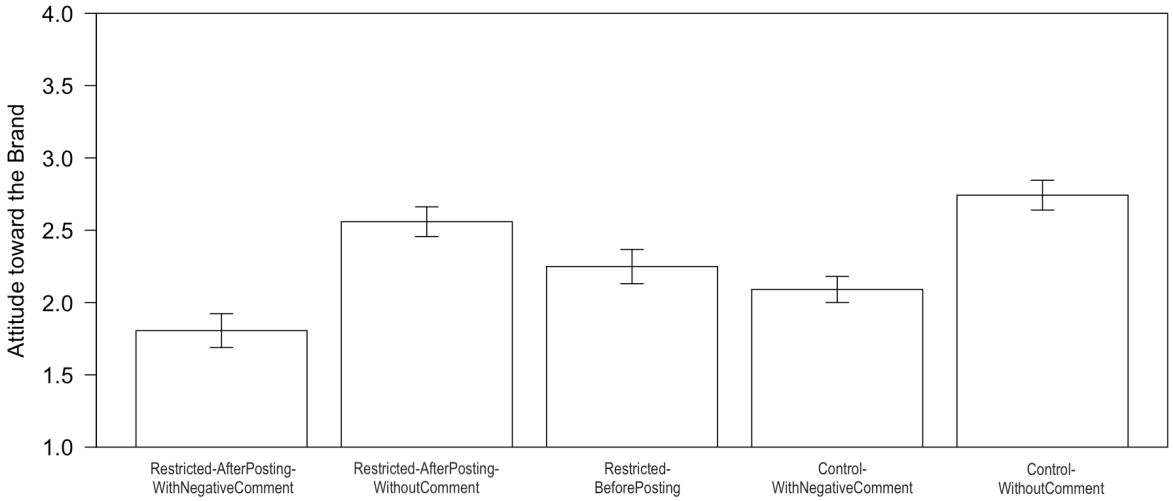


Fig. 11. The Effect of the Different Conditions on Suspicion

11.4. Discussion

This study demonstrates that the adverse consequences of constraining the ability to respond to social media posts on a brand outweigh the negative effects from the received comments. However, if we can anticipate negative comments, such as those related to a post discussing a company controversy, it is preferable to limit commenting before posting rather than after receiving such comments.

12. General discussion

This paper scrutinized the phenomenon of limiting the comments on social media, which pertains to the scenario where a user restricts or disables the ability to comment, preventing others' commenting. Through a series of seven studies—comprising one field study and six online experiments—the research unveiled that limiting the comments holds a significant impact on suspicion, ultimately leading to a decrease in attitude toward the brand for users who employ this limitation. Furthermore, the studies underscored that this effect is attributed to consumers interpreting the restriction of comments as a cue of information censorship.

The findings indicate that despite the introduction of the option to restrict the ability to comment with the intention of safeguarding privacy—a practical and valuable consideration in the current internet landscape—its implementation in real-life scenarios could potentially jeopardize users' perceptions of both the post and its author. The prevalence of extensive bots on social media, fabricated reviews and comments, deceptive online advertisements, and instances of cyberbullying have rendered individuals more vigilant and skeptical toward online information. Consequently, the strategy of limiting the comments appears to fall prey to this overarching suspicion. The results of this paper underscore that even when a legitimate rationale is present and effectively communicated to users, it fails to counteract the adverse impact of limiting the comments on users' attitudes.

More interestingly, our findings highlight a significant insight: even in instances where there is no discernible rationale for information censorship—such as when sharing content of low importance or with no possible hidden information—individuals still harbor suspicion and negativity toward the post. This phenomenon could be attributed to the notion that encountering a post with a restricted comment, even if no censorship is actually warranted, triggers a perception of censorship. This perception stems from the fundamental role of the comment section as a means of communication for each post. Consequently, the exclusion of this feature for seemingly random posts might evoke a sense of diminished freedom of speech, contributing to the observed suspicion and negative sentiments.

In conclusion, it is crucial to emphasize that although a majority of our studies incorporated manipulations resembling Twitter, substantial evidence suggests that the documented effects possess broader applicability and extend to various other social media platforms. For instance, Study 4 introduced a self-designed social media platform to demonstrate the generalizability of observed effects across different platforms, so long as a visual cue indicating the limitation of the comments on the post is present. This visual cue is

prevalent across all contemporary social media platforms. Furthermore, our exploration of diverse content types and industries underscores the resilience of our model concerning the nature of the post's content.

12.1. Theoretical contributions

The phenomenon of consumer suspicion, as elucidated by Buller and Burgoon (1996), revolves around the inherent uncertainty tied to the authenticity of a sender's message. This overarching sense of skepticism, fueled by the ongoing ambiguity stemming from perceived deception, creates an atmosphere where the trustworthiness of presented information is subject to doubt. The unregulated terrain of today's internet landscape has amplified public awareness of potential deceptive practices on social media platforms. Whether through the dissemination of brand propaganda via automated bots or the prevalence of misleading advertisements that drastically deviate from actual product realities (Hendy, 2019), individuals find themselves increasingly exposed to an expanding spectrum of deceitful content.

Furthermore, individuals in the modern era grapple with an overwhelming influx of information on a daily basis. Engaging in exhaustive rational analysis to grasp and process this torrent of data proves to be an impractical endeavor. The volume of information encountered through social media continually swells. As Simon (1971. P.40) astutely observed, information demands the attention of its recipients, thereby ushering in a "wealth of information creates a poverty of attention" scenario. The rise of social media platforms has further exacerbated the information overload experienced by consumers (Rodriguez et al., 2014).

In light of this predicament, consumers are compelled to allocate their attention judiciously among a multitude of information sources (Simon, 1955), invariably necessitating a trade-off between efficiency and accuracy when assessing the credibility of information

(Fogg, 2003). As a response, consumers often resort to specific cues and cognitive shortcuts that enable them to make decisions (Tversky & Kahneman, 1974) and discern the credibility of content on social media (Hilligoss & Rieh, 2008).

For instance, individuals gauge a user's authenticity based on factors such as their profile picture (Xu, 2014), the presence of a verified badge, and even linguistic aspects within their posts, such as grammar and punctuation (Morris et al., 2012). Additionally, people might delve into intricate social cues, such as a source's follower-to-following ratio (Valsesia et al., 2020) or the ratio of comments to likes on a post (Rezaee et al., TBD), in their quest to gain insights into content credibility. In summary, the escalating climate of consumer skepticism within the realm of social media directly stems from the evolving landscape characterized by pervasive deception and information overload. The conscious reliance on distinct social cues for decision-making underscores consumers' strategic adaptation to navigate this intricate digital environment.

Accordingly, this paper demonstrates that consumers perceive the restriction of the comments as a potent cue for potential deceptive information censorship. Irrespective of the presence of valid and legitimate reasons for limiting comments, our studies reveal that the mere act itself evokes a sense of censorship among users, even when there might be no cause for concern. Our findings support the hypothesis that encountering a post with restricted comments fosters increased suspicion toward the content, and individuals may respond by seeking additional information, disseminating negative word-of-mouth, and generally harboring negative attitudes toward the brand as a whole.

Furthermore, this paper thoroughly investigates the mediating pathway between limiting the comments and the resulting attitude toward the brand, with the goal of identifying the central mechanism underlying this phenomenon. Study 2 demonstrates that the effect indeed arises from perceived censorship, dispelling other theoretically plausible explanations

in the literature. These explanations encompass concerns about customer-responsiveness, negative reactions to content due to inconsistency with the actions of others, and even potential positive brand attitudes stemming from the acknowledgment of the brand's resilience against cyberbullying attacks.

12.2. Practical contributions

The insights presented in this article offer direct applicability to practitioners who leverage social media as a communication platform for their companies or products. Firstly, the findings underscore the importance of maintaining an open two-way communication channel, regardless of the validity of the reasons for limiting the comments. The research highlights that a considerable number of consumers heavily rely on comments as a primary source for decision-making and engaging with brands. Consequently, even in scenarios where there is no intent to censor or withhold information, the act of limiting the comments might be interpreted as constraining users' freedom of expression. This can potentially lead to a decline in brand attitude. Study 5, for instance, provides concrete evidence that even when confronted with hate speech or trolling, restriction of the ability to comment can adversely impact brand perception, despite explanations being provided to users.

Furthermore, certain brands and influencers opt for a consistent strategy of restricting comments across their social media posts. In such cases, one could argue that users might perceive this as a standard mode of communication for the brand, rather than an indicator of censorship. However, Study 4 refutes this assumption by demonstrating that even when the limitation is consistently applied, it still contributes to a decrease in users' brand attitudes.

Study 6 provides a crucial insight for marketers: when it comes to managing your brand's reputation on social media, it is essential to carefully weigh the consequences of restricting user engagement. Interestingly, the research demonstrates that the detrimental

impact of limiting responses can be more significant than the potential harm caused by negative feedback itself.

What is particularly intriguing is that this study offers a strategic approach to address this challenge. If you possess the ability to foresee the likelihood of negative comments, especially in cases where your post delves into a controversial issue related to your company, the most prudent course of action is to proactively restrict commenting before the post is made. In other words, exercising control before the content is published can be more effective in safeguarding your brand's online image than waiting to react after encountering negative comments.

In essence, these findings suggest that regardless of the context or reasoning, limiting the comments can inadvertently signal suspicion and curtailment of communication. The practical implication is clear: brands are advised to exercise caution when considering the restriction of comments, even for posts that may seem less significant or for consistency purposes. Open dialogue and engagement on social media platforms remain pivotal to fostering positive brand perceptions and mitigating potential skepticism among consumers.

In a nutshell, the central contribution of this paper can be distilled into a succinct guideline: Do not limit the comments!

12.3. Limitations and future research

Indeed, exploring the information-seeking behavior that individuals might undertake after encountering a post with restricted comments is a valuable avenue for future research. Understanding how users respond to this perceived censorship by actively seeking out information that might have been hidden could shed light on the intricate dynamics between restriction of ability to comment, suspicion, and information retrieval. Investigating whether individuals, driven by reactance theory or curiosity, are prompted to focus deeper into other sources or platforms to uncover potentially censored information could provide a more

comprehensive view of the consequences of limiting comments. This aspect not only adds depth to the understanding of the phenomenon but also offers practical insights for brands and content creators aiming to maintain transparency and retain consumer trust in an environment increasingly defined by skepticism.

Moreover, if we broaden our perspective to consider the act of blocking a user as a manifestation of online censorship, this phenomenon emerges as a valuable area of investigation. Such an exploration is particularly pertinent due to its potential to generate actively negative attitudes and lead to the dissemination of negative electronic Word-of-Mouth (eWOM) concerning brands. In essence, studying the impact of user blocking as a form of online censorship on brand reputation could provide critical insights into how it influences consumer behavior and brand perception.

Furthermore, certain brands and influencers opt to disable the comments as a strategic move to enhance their content's virality. This strategy rests on the belief that by restricting comments, individuals seeking to express their opinions are compelled to resort to sharing the content (or quote-retweeting it) to effectively convey their sentiments. This is thought to leverage the greater visibility that sharing can offer compared to simple commenting. However, it is now pertinent to explore the implications for individuals who hold a strong affinity for a brand or product and habitually leave positive comments. How might their behavior shift if they encounter restricted comments? Could this lead to an increase in content sharing, or does it potentially transform their positive feedback into negative sentiment toward the brand due to the limitation on comments? This investigation could unveil the delicate balance between encouraging user engagement and preserving a positive brand perception amid comment limitations.

Conclusion

This dissertation comprises three distinct essays, all centered on the intricate dynamics of decision-making processes within the realm of social media. It explores the multifaceted aspects of how individuals employ various heuristics, cues, and signals to navigate the vast landscape of social media platforms when making decisions, whether it be for staying informed about current news, patronizing their preferred brands, or following celebrities and influencers.

The impetus for this investigation is rooted in the ever-expanding digital age. In this era, an overabundance of information inundates users, leading to a noticeable decline in the inclination and motivation for engaging in central, systematic thinking, as pointed out by Pee (2012). This dissertation strives to shed light on the cognitive shortcuts and mechanisms that individuals rely on when navigating the complex social media environment, ultimately contributing to our understanding of decision-making processes in the digital era.

The first article conducts a systematic review of literature centered on decision-making within digital environments, summarizing the research from business, management, and communication journals spanning the last decade. The primary focus of this review was on the cues and signals that individuals rely on when making decisions. These cues encompass the assessment of user credibility, the authenticity of social media posts, and the sponsored content generated by influencers across various platforms, including Twitter, Instagram, and YouTube. In conclusion, this essay categorized the literature and identified gaps in the existing research, and provides practical recommendations for brands and influencers to better align their content creation with the decision-making processes of their target users and consumers.

Addressing a crucial research gap identified in the first article, the second essay centers on the assessment of social media post credibility, specifically tweets, when users lack

access to the source of information and rely solely on available post statistics, such as the number of comments and likes. This research explores the concept of 'ratioing' and its effect of the perceived credibility of social media posts. The hypothesis tested in this study posits that users consider high-ratio posts to be less credible than low-ratio ones. This hypothesis was investigated through a field study and five experimental studies. The findings suggest that high ratios are perceived as a lack of consensus, which, in turn, impacts users' purchasing behavior and intentions. This article contributes to our comprehension of information processing on social media and provides insights valuable to companies. It offers a content creation strategy aimed at mitigating the adverse effects of ratioing, particularly in the case of potentially controversial posts.

The third article addresses another research gap highlighted in the first essay, focusing on how consumers react to the comment sections of posts, particularly when the user who made the post restricts comments from others, including their followers. The hypothesis posits that when social media posts restrict others from commenting, it may adversely affect consumers' perceptions of the brand associated with the post. Specifically, consumers may interpret the inability to comment as an infringement on free speech, leading to unfavorable opinions about the brand or individual responsible for the post. To validate this hypothesis, a field study and seven online experiments were conducted. The results demonstrate that closing the comment section influences attitudes toward the brand due to perceived censorship, resulting in increased user suspicion. Furthermore, this essay shows that receiving negative comments has less of a negative effect on users' attitude toward the brand than restricting the ability to comment after receiving such negative feedback.

In conclusion, it is abundantly clear that marketers need to maintain a high level of vigilance when it comes to the cues and signals present in the online communication platforms they utilize. Even seemingly minor or inconspicuous cues can wield significant

influence over their target audiences, and neglecting to study and comprehend these cues can lead to unexpected and adverse consequences. Therefore, gaining a deep understanding of the fundamental dynamics of these cues is paramount for successful online marketing strategies.

Furthermore, marketers can harness these cues, such as the comment-to-like ratio, as invaluable social media metrics to continuously gauge and manage the audience's perception of their posts. This metric offers a real-time and actionable insight into how well their content is resonating with their audience. By carefully monitoring and responding to changes in these cues, marketers can adapt and refine their strategies to ensure that they maintain a positive and engaging online presence. In the ever-evolving landscape of digital marketing, staying attuned to these cues is not just a choice, but a necessity for those looking to thrive and succeed in the online realm.

References

Aditya, R. N. (2001). The psychology of deception in marketing: A conceptual framework for research and practice. *Psychology & Marketing*, *18*(7), 735-761. https://doi.org/10.1002/mar.1028

Agnihotri, R., Dingus, R., Hu, M. Y., & Krush, M. T. (2016). Social media: Influencing customer satisfaction in B2B sales. *Industrial Marketing Management*, *53*, 172-180. https://doi.org/10.1016/j.indmarman.2015.09.003

Ahluwalia, R., & Burnkrant, R. E. (2004). Answering questions about questions: A persuasion knowledge perspective for understanding the effects of rhetorical questions. *Journal of Consumer Research, 31*(1), 26–42. https://doi.org/10.1086/383421

Algi, A. (2018). Consumer trust and intention to buy in Indonesia Instagram stores. In *2018 3rd International Conference on Information Technology, Information System and Electrical Engineering (ICITISEE)* (pp. 199-203). IEEE. https://doi.org/10.1109/ICITISEE.2018.8721033

Almgren, S. M., & Olsson, T. (2016). Commenting, sharing and tweeting news. *Nordicom Review*, *37*(2), 67. https://doi.org/10.1515/nor-2016-0018

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596. https://doi.org/10.1037/0022-3514.63.4.596

Avram, M., Micallef, N., Patil, S., & Menczer, F. (2020). Exposure to Social Engagement Metrics Increases Vulnerability to Misinformation. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-033

Balaji, M. S., Jiang, Y., & Jha, S. (2021). Nanoinfluencer marketing: How message features affect credibility and behavioral intentions. *Journal of Business Research*, *136*, 293-304. https://doi.org/10.1016/j.jbusres.2021.07.049

Bart, Y., Shankar, V., Sultan, F., & Urban, G. L. (2005). Are the drivers and role of online trust the same for all web sites and consumers? A large-scale exploratory empirical

study. *Journal of Marketing, 69*(4), 133–152. https://doi.org/10.1509/jmkg.2005.69.4.133

Beneke, J., de Sousa, S., Mbuyu, M., & Wickham, B. (2016). The effect of negative online customer reviews on brand equity and purchase intention of consumer electronics in South Africa. *The International Review of Retail, Distribution and Consumer Research*, *26*(2), 171-201. https://doi.org/10.1080/09593969.2015.1068828

Bitter, S., & Grabner-Kräuter, S. (2016). Consequences of customer engagement behavior: when negative Facebook posts have positive effects. *Electronic Markets*, *26*(3), 219-231. https://doi.org/10.1007/s12525-016-0220-7

Bolton, M. (2004). Customer centric business processing. *International Journal of Productivity and Performance Management*, *53*(1), 44-51. https://doi.org/10.1108/17410400410509950

Boutet, I., LeBlanc, M., Chamberland, J. A., & Collin, C. A. (2021). Emojis influence emotional communication, social attributions, and information processing. *Computers in Human Behavior*, *119*, 106722. https://doi.org/10.1016/j.chb.2021.106722

Bradley, P. (2010). Be where the conversations are: The critical importance of social media. *Business Information Review*, *27*(4), 248-252. https://doi.org/10.1177/0266382110390976

Brehm, J. W. (1966). A Theory of Psychological Reactance. *Academic, Oxford*.

Brodie, R. J., Ilic, A., Juric, B., & Hollebeek, L. (2013). Consumer engagement in a virtual brand community: An exploratory analysis. *Journal of Business Research*, *66*(1), 105-114. https://doi.org/10.1016/j.jbusres.2011.07.029

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal deception theory. *Communication Theory, 6*(3), 203-242. https://doi.org/10.1111/j.1468-2885.1996.tb00127.x

Burger, J. M., Horita, M., Kinoshita, L., Roberts, K., & Vera, C. (1997). Effects on time on the norm of reciprocity. *Basic and Applied Social Psychology*, *19*(1), 91-100. https://doi.org/10.1207/s15324834basp1901_7

Burgoon, J. K., Buller, D. B., Ebesu, A. S., White, C. H., & Rockwell, P. A. (1996). Testing interpersonal deception theory: Effects of suspicion on communication behaviors and

perceptions. *Communication Theory*, *6*(3), 243-267. https://doi.org/10.1111/j.1468-2885.1996.tb00128.x

Camacho, S., Hassanein, K., & Head, M. (2014). Understanding the factors that influence the perceived severity of cyber-bullying. In *HCI in Business: First International Conference, HCIB 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings 1* (pp. 133-144). Springer International Publishing. https://doi.org/10.1007/978-3-319-07293-7_13

Carr, C. T., & Hayes, R. A. (2015). Social media: Defining, developing, and divining. *Atlantic Journal of Communication*, *23*(1), 46-65. https://doi.org/10.1080/15456870.2015.972282

Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on Twitter. *Proceedings of the 20th International Conference on World Wide Web*, 675–684. https://doi.org/10.1145/1963405.1963500

Chaiken, S. (1987). The heuristic model of persuasion. *In Social Influence: The Ontario Symposium* (Vol. 5, pp. 3-39).

Chaiken, S., & Stangor, C. (1987). Attitudes and attitude change. *Annual Review of Psychology*. https://doi.org/10.1146/annurev.ps.38.020187.003043

Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context.

Chen, P. Y., Wu, S. Y., & Yoon, J. (2004). The impact of online recommendations and consumer feedback on sales. https://aisel.aisnet.org/icis2004/58

Counts, S., & Fisher, K. (2011). Taking it all in? visual attention in microblog consumption. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 5, No. 1, pp. 97-104). https://doi.org/10.1609/icwsm.v5i1.14103

De Maeyer, P. (2012). Impact of online consumer reviews on sales and price strategies: A review and directions for future research. *Journal of Product & Brand Management*, *21*(2), 132-139. https://doi.org/10.1108/10610421211215599

De Vries, E. L. (2019). When more likes is not better: the consequences of high and low likes-to-followers ratios for perceived account credibility and social media marketing

effectiveness. *Marketing Letters*, *30*(3-4), 275-291. https://doi.org/10.1007/s11002-019-09496-6

DeCarlo, T. E. (2005). The effects of sales message and suspicion of ulterior motives on salesperson evaluation. *Journal of Consumer Psychology*, *15*(3), 238-249. https://doi.org/10.1207/s15327663jcp1503_9

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, *49*(10), 1407-1424. https://doi.org/10.1287/mnsc.49.10.1407.17308

DePaulo, P. J., & DePaulo, B. M. (1989). Can deception by salespersons and customers be detected through nonverbal behavioral cues? *Journal of Applied Social Psychology*, *19*(18), 1552-1577. https://doi.org/10.1111/j.1559-1816.1989.tb01463.x

Eckstein, H. (1997). Congruence theory explained.

Eisend, M. (2006). Source credibility dimensions in marketing communication–A generalized solution. *Journal of Empirical Generalisations in Marketing Science*, *10*(2).

Erb, H.-P., Bohner, G., Hewstone, M., Werth, L., & Reinhard, M.-A. (2006). Large Minorities and Small Majorities: Interactive Effects of Inferred and Explicit Consensus on Attitudes. *Basic and Applied Social Psychology*, *28*(3), 221–231. https://doi.org/10.1207/s15324834basp2803_2

Ewing, M. T., Salzberger, T., & Sinkovics, R. R. (2005). An alternate approach to assessing cross-cultural measurement equivalence in advertising research. *Journal of Advertising*, *34*(1), 17-36. https://doi.org/10.1080/00913367.2005.10639181

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, *68*(6), 1261-1270. https://doi.org/10.1016/j.jbusres.2014.11.006

Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of Internet information credibility. *Journalism & Mass Communication Quarterly*, *77*(3), 515-540. https://doi.org/10.1177/107769900007700304

Flanagin, A. J., Metzger, M. J., Pure, R., Markov, A., & Hartsell, E. (2014). Mitigating risk in ecommerce transactions: Perceptions of information credibility and the role of user-

generated ratings in product quality and purchase intention. *Electronic Commerce Research*, *14*(1), 1–23. https://doi.org/10.1007/s10660-014-9139-2

Floh, A., Koller, M., & Zauner, A. (2009). The impact of perceived valence, perceived information credibility and valence intensity of online reviews on purchase intentions. https://aisel.aisnet.org/iceb2009/19

Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, 722–723. https://doi.org/10.1145/765891.765951

Freshwater, H. (2004). Towards a redefinition of censorship. In *Censorship & Cultural Regulation in the Modern Age* (pp. 217-237). Brill. https://doi.org/10.1163/9789401200950_010

Frias, D. M., Rodriguez, M. A., & Castaneda, J. A. (2008). Internet vs. travel agencies on pre-visit destination image formation: an information processing view. *Tourism Management*, 29, 163-179. https://doi.org/10.1016/j.tourman.2007.02.020

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research, 21*(1), 1. https://doi.org/10.1086/209380

Gaziano, C., & McGrath, K. (1986). Measuring the concept of credibility. *Journalism Quarterly*, *63*(3), 451-462. https://doi.org/10.1177/107769908606300301

Ghose, A., & Ipeirotis, P. G. (2008). Estimating the socio-economic impact of product reviews: Mining text and reviewer characteristics.

Goh, K. Y., Heng, C. S., & Lin, Z. (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user-and marketer-generated content. *Information Systems Research*, *24*(1), 88-107. https://doi.org/10.1287/isre.1120.0469

Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. *In Simple heuristics that make us smart* (pp. 37-58). Oxford University Press.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: the recognition heuristic. *Psychological Review*, *109*(1), 75. https://doi.org/10.1037/0033-295X.109.1.75

Gross, B. M. (1964). The managing of organizations: The administrative struggle (Vol. 2). *[New York]: Free Press of Glencoe*.

Habel, J., Kassemeier, R., Alavi, S., Haaf, P., Schmitz, C., & Wieseke, J. (2020). When do customers perceive customer centricity? The role of a firm's and salespeople's customer orientation. *Journal of Personal Selling & Sales Management*, *40*(1), 25-42. https://doi.org/10.1080/08853134.2019.1631174

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling.

Hendy, N. (2019), "The Sidney Morning Herald, 'negative review? Now you can just delete it'", available at: www.smh. com.au/business/small-business/negative-review-now-you-can- just-delete-it-20190613-p51x8q.html

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, *18*(1), 38-52. https://doi.org/10.1002/dir.10073

Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, *44*(4), 1467–1484. https://doi.org/10.1016/j.ipm.2007.10.001

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, *15*(4), 635-650. https://doi.org/10.1086/266350

Jang, S., Prasad, A., & Ratchford, B. T. (2017). Consumer search of multiple information sources and its impact on consumer price satisfaction. *Journal of Interactive Marketing, 40*(1), 24-40. https://doi.org/10.1016/j.intmar.2017.06.004

Johnson, C. L. (2011). An examination of the primary and secondary effects of cyber-bullying: Development and testing of a cyber-bullying moderator/mediator model. *Wayne State University*.

Kalbfleisch, P. J. (2003). Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. *Communication Yearbook 27*, 307-350. https://doi.org/10.1080/23808985.2003.11679029

Ki, C. W. C., Cuevas, L. M., Chong, S. M., & Lim, H. (2020). Influencer marketing: Social media influencers as human brands attaching to followers and yielding positive marketing results by fulfilling needs. *Journal of Retailing and Consumer Services*, *55*, 102133. https://doi.org/10.1016/j.jretconser.2020.102133

Kim, J.-H., & Song, H. (2020). The influence of perceived credibility on purchase intention via competence and authenticity. *International Journal of Hospitality Management*, *90*, 102617. https://doi.org/10.1016/j.ijhm.2020.102617

Kiousis, S. (2006). Exploring the Impact of Modality on Perceptions of Credibility for Online News Stories. *Journalism Studies*, *7*(2), 348–359. https://doi.org/10.1080/14616700500533668

Könsgen, R., Schaarschmidt, M., Ivens, S., & Munzel, A. (2018). Finding meaning in contradiction on employee review sites—Effects of discrepant online reviews on job application intentions. *Journal of Interactive Marketing*, *43*, 165-177. https://doi.org/10.1016/j.intmar.2018.05.001

Le, T. V., ALANG, T., & TRAN, Q. T. (2021). How YouTube Influencers Impact Customers' Purchase Intention: An Empirical Study of Cosmetic Brands in Vietnam. *The Journal of Asian Finance, Economics and Business*, *8*(9), 101-111. http://dx.doi.org/10.13106/jafeb.2021.vol8.no9.0101

Lee, J. Y. (2015). The Number of Likes Associated with Given Health-Related Messages on Facebook: The Moderating Effect of Value Involvement [Ph.D.]. https://www.proquest.com/docview/1718185168/abstract/C5A75CE07AA747D4PQ/1

Lee, J. Y., & Sundar, S. S. (2013). To tweet or to retweet? That is the question for health professionals on Twitter. *Health Communication*, *28*(5), 509-524. http://dx.doi.org/10.1080/10410236.2012.700391.

Lee, Y. H., & Mason, C. (1999). Responses to information incongruency in advertising: The role of expectancy, relevancy, and humor. *Journal of Consumer Research, 26*(2), 156-169. https://doi.org/10.1086/209557

Levordashka, A., Utz, S., & Ambros, R. (2016, March). What's in a like? Motivations for pressing the like button. *In the Tenth International AAAI Conference on Web and Social Media*. https://doi.org/10.1609/icwsm.v10i1.14768

Li, R., & Sundar, S. S. (2022). Can interactive media attenuate psychological reactance to health messages? A study of the role played by user commenting and audience metrics in persuasion. *Health Communication*, *37*(11), 1355-1367. https://doi.org/10.1080/10410236.2021.1888450

Lin, X. (2016). Social Media And Credibility Indicator: The Effects Of Bandwagon And Identity Cues Within Online Health And Risk Contexts. *Theses and Dissertations–Communication*. http://dx.doi.org/10.13023/ETD.2016.059

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management*, *47*, 140-151. https://doi.org/10.1016/j.tourman.2014.09.020

Lucassen, T., & Schraagen, J. M. (2013). The influence of source cues and topic familiarity on credibility evaluation. *Computers in Human Behavior*, *29*(4), 1387-1392. https://doi.org/10.1016/j.chb.2013.01.036

Mackie, D. M., Worth, L. T., & Asuncion, A. G. (1990). Processing of persuasive in-group messages. *Journal of Personality and Social Psychology*, *58*(5), 812. https://doi.org/10.1037/0022-3514.58.5.812

MacInnis, D. J., Rao, A. G., & Weiss, A. M. (2002). Assessing when increased media weight of real-world advertisements helps sales. *Journal of Marketing Research*, *39*(4), 391-407. https://doi.org/10.1509/jmkr.39.4.391.19

Maheswaran, D., & Chaiken, S. (1991). Promoting systematic processing in low-motivation settings: Effect of incongruent information on processing and judgment. *Journal of Personality and Social Psychology*, *61*(1), 13. https://psycnet.apa.org/doi/10.1037/0022-3514.61.1.13

Martínez-López, F. J., Anaya-Sánchez, R., Fernández Giordano, M., & Lopez-Lopez, D. (2020). Behind influencer marketing: key marketing decisions and their effects on followers' responses. *Journal of Marketing Management*, *36*(7-8), 579-607. https://doi.org/10.1080/0267257X.2020.1738525

McDaniel, B. (2021). Popular music reaction videos: Reactivity, creator labor, and the performance of listening online. *New Media & Society*, *23*(6), 1624-1641. https://doi.org/10.1177/1461444820918549

Meinert, J., & Krämer, N. C. (2022). How the expertise heuristic accelerates decision-making and credibility judgments in social media by means of effort reduction. *Plos one*, *17*(3), e0264428. https://doi.org/10.1371/journal.pone.0264428

Metaxas, P., Mustafaraj, E., Wong, K., Zeng, L., O'Keefe, M., & Finn, S. (2015). What do retweets indicate? Results from user survey and meta-review of research. In *Proceedings of the International AAAI Conference on Web and Social Media*(Vol. 9, No. 1, pp. 658-661). https://doi.org/10.1609/icwsm.v9i1.14661

Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & McCann, R. M. (2003). Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. *Annals of the International Communication Association*, *27*(1), 293-335. https://doi.org/10.1080/23808985.2003.11679029

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, *60*(3), 413-439. https://doi.org/10.1111/j.1460-2466.2010.01488.x

Minot, J. R., Arnold, M. V., Alshaabi, T., Danforth, C. M., & Dodds, P. S. (2021). Ratioing the President: An exploration of public engagement with Obama and Trump on Twitter. *Plos One*, *16*(4), e0248880. https://doi.org/10.1371/journal.pone.0248880

Mohan, B., Schlager, T., Deshpandé, R., & Norton, M. I. (2018). Consumers avoid buying from firms with higher CEO-to-worker pay ratios. *Journal of Consumer Psychology*, *28*(2), 344-352. https://doi.org/10.1002/jcpy.1033

Morris, M. R., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012). Tweeting is believing? Understanding microblog credibility perceptions. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 441–450. https://doi.org/10.1145/2145204.2145274

Moon, S., Kim, M. Y., & Bergey, P. K. (2019). Estimating deception in consumer reviews based on extreme terms: Comparison analysis of open vs. closed hotel reservation platforms. *Journal of Business Research*, *102*, 83-96. https://doi.org/10.1016/j.jbusres.2019.05.016

Muda, M., & Hamzah, M. I. (2021). Should I suggest this YouTube clip? The impact of UGC source credibility on eWOM and purchase intention. *Journal of Research in Interactive Marketing*, *15*(3), 441-459. https://doi.org/10.1108/JRIM-04-2020-0072

Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon. com. *MIS Quarterly*, 185-200. https://doi.org/10.2307/20721420

Muntinga, D. G., Moorman, M., & Smit, E. G. (2011). Introducing COBRAs. *International Journal of Advertising*, *30*(1), 13–46. https://doi.org/10.2501/IJA-30-1-013-046

O'Keefe, D. J. (1990). Persuasion: Theory and practice. *Newbury Park*: Sage.

Obermiller, C., & Spangenberg, E. R. (1998). Development of a scale to measure consumer skepticism toward advertising. *Journal of Consumer Psychology*, *7*(2), 159-186. https://doi.org/10.1207/s15327663jcp0702_03

Oh, J. Y., & Aukerman, R. A. (2013). Freedom of speech and censorship in the Internet. *International Journal of Management & Information Systems (IJMIS)*, *17*(4), 251-256. https://doi.org/10.19030/ijmis.v17i4.8101

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological Review, 62*(1), 42. https://doi.org/10.1037/h0048153

Ott, M., Cardie, C., & Hancock, J. T. (2013, June). Negative deceptive opinion spam. In *Proceedings of the 2013 conference of the North American chapter of the Association for computational linguistics: Human Language Technologies* (pp. 497-501).

Oza, S. S., Srivastava, J., & Koukova, N. T. (2010). How suspicion mitigates the effect of influence tactics. *Organizational Behavior and Human Decision Processes, 112*(1), 1–10. https://doi.org/10. 1016/j.obhdp.2009.11.006

Özbölük, T., & Akdoğan, K. (2022). The role of online source credibility and influencer identification on consumers' purchase decisions. *International Journal of Internet Marketing and Advertising*, *16*(1-2), 165-185. https://doi.org/10.1504/IJIMA.2022.120974

Paivio, A. (2013). Imagery and verbal processes. *Psychology Press*.

Park, D. H., & Lee, J. (2008). eWOM overload and its effect on consumer behavioral intention depending on consumer involvement. *Electronic Commerce Research and Applications*, *7*(4), 386-398. https://doi.org/10.1016/j.elerap.2007.11.004

Pee, L. G. (2012). Trust of information on social media: an elaboration likelihood model.

Permana, I. P. H., & Meinarni, N. P. S. (2021). Ratio analysis on TikTok (social media) for qualitative research using explorative methods. *Jurnal Ekonomi & Bisnis JAGADITHA*, *8*(1), 30-38. https://doi.org/10.22225/jj.8.1.2944.30-38

Petty, R. E. (2013). Two routes to persuasion: State of the art. *International Perspectives on Psychological Science*, *2*(1), 229-247.

Petty, R. E. (2018). Attitudes and persuasion: Classic and Contemporary Approaches. *Routledge.*

Petty, R. E., Cacioppo, J. T., Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion (pp. 1-24). *Springer New York*.

Plotkina, D., Munzel, A., & Pallud, J. (2020). Illusions of truth— Experimental insights into human and algorithmic detections of fake online reviews. *Journal of Business Research, 109*, 511– 523. https://doi.org/10.1016/j.jbusres.2018.12.009

Ranganathan, C. (2012). The role of extrinsic cues in consumer decision process in online shopping environments. *Journal of Electronic Commerce in Organizations (JECO)*, *10*(1), 52-71.

Reczek, R. W., Trudel, R., & White, K. (2018). Focusing on the forest or the trees: How abstract versus concrete construal level predicts responses to eco-friendly products. *Journal of Environmental Psychology*, *57*, 87-98. https://doi.org/10.1016/j.jenvp.2018.06.003

Regan, D. T. (1971). Effects of a favor and liking on compliance. *Journal of Experimental Social Psychology*, *7*(6), 627-639. https://doi.org/10.1016/0022-1031(71)90025-4

Riedl, M. J., Naab, T. K., Masullo, G. M., Jost, P., & Ziegele, M. (2021). Who is responsible for interventions against problematic comments? Comparing user attitudes in Germany and the United States. *Policy & Internet*, *13*(3), 433-451. https://doi.org/10.1002/poi3.257

Riedl, R., Mohr, P. N., Kenning, P. H., Davis, F. D., & Heekeren, H. R. (2014). Trusting humans and avatars: A brain imaging study based on evolution theory. *Journal of Management Information Systems*, *30*(4), 83-114. https://doi.org/10.2753/MIS0742-1222300404

Rieskamp, J., & Hoffrage, U. (1999). When do people use simple heuristics, and how can we tell?

Rodriguez, M. G., Gummadi, K., & Schoelkopf, B. (2014, May). Quantifying information overload in social media and its impact on social contagions. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 8, No. 1, pp. 170-179). https://doi.org/10.1609/icwsm.v8i1.14549

Royo-Vela, M., & Casamassima, P. (2011). The influence of belonging to virtual brand communities on consumers' affective commitment, satisfaction, and word-of-mouth advertising: The ZARA case. *Online Information Review*, *35*(4), 517-542. https://doi.org/10.1108/14684521111161918

media-and-tourism-industry-statistics.

Sah, S., Malaviya, P., & Thompson, D. (2018). Conflict of interest disclosure as an expertise cue: Differential effects due to automatic versus deliberative processing. *Organizational Behavior and Human Decision Processes*, *147*, 127-146. https://doi.org/10.1016/j.obhdp.2018.05.008

Santos, S. (2014). 2012e2013 Social media and tourism industry statistics. *Stikky Media*. from http://www.stikkymedia.com/blog/2012-2013-social-

Sashi, C. M. (2012). Customer engagement, buyer-seller relationships, and social media. *Management Decision*, *50*(2), 253-272. https://doi.org/10.1108/00251741211203551

Šerić, M., Ozretić-Došen, Đ., & Škare, V. (2020). How can perceived consistency in marketing communications influence customer–brand relationship outcomes? *European Management Journal*, *38*(2), 335-343. https://doi.org/10.1016/j.emj.2019.08.011

Shah, D., Rust, R. T., Parasuraman, A., Staelin, R., & Day, G. S. (2006). The path to customer centricity. *Journal of Service Research*, *9*(2), 113-124. https://doi.org/10.1177/1094670506294666

Shugan, S. M. (2006). Who is afraid to give freedom of speech to marketing folks? *Marketing Science*, *25*(5), 403-410. https://doi.org/10.1287/mksc.1060.0236

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99–118. https://doi.org/10.2307/1884852

Simon, H. A. (1971). Designing organizations for an information-rich world. *Computers, Communications, and the Public Interest*, *72*-37.

Soh, H., Reid, L. N., & King, K. W. (2009). Measuring Trust in Advertising. *Journal of Advertising*, *38*(2), 83–104. https://doi.org/10.2753/JOA0091-3367380206

Stöckli, S., & Hofer, D. (2020). Susceptibility to social influence predicts behavior on Facebook. *PloS One*, *15*(3). https://doi.org/10.1371/journal.pone.0229337

Stubb, C., & Colliander, J. (2019). "This is not sponsored content"–The effects of impartiality disclosure and e-commerce landing pages on consumer responses to social media

influencer posts. *Computers in Human Behavior*, *98*, 210-222. https://doi.org/10.1016/j.chb.2019.04.024

Sundar, S. S. (2008). The MAIN model: A heuristic approach to understanding technology effects on credibility (pp. 73-100). *Cambridge, MA: MacArthur Foundation Digital Media and Learning Initiative.*

Timoshenko, A., & Hauser, J. R. (2019). Identifying customer needs from user-generated content. *Marketing Science*, *38*(1), 1-20. https://doi.org/10.1287/mksc.2018.1123

Todorov, A., Chaiken, S., & Henderson, M. D. (2002). The heuristic-systematic model of social information processing. *The Persuasion Handbook: Developments in Theory and Practice*, *23*, 195-211.

Tsiakali, K. (2018). User-generated content versus marketing-generated-content: Personality and content influence on traveler's behavior. *Journal of Hospitality Marketing & Management*, *27*(8), 946-972. https://doi.org/10.1080/19368623.2018.1477643

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124

Vaidya, T., Votipka, D., Mazurek, M. L., & Sherr, M. (2019, May). Does being verified make you more credible? Account verification's effect on tweet credibility. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). https://doi.org/10.1145/3290605.3300755

Valsesia, F., Proserpio, D., & Nunes, J. C. (2020). The positive effect of not following others on social media. *Journal of Marketing Research*, *57*(6), 1152-1168. https://doi.org/10.1177/0022243720915467

Vareberg, K. R., Vogt, O., & Berndt, M. (2023). Putting your best face forward: How instructor emoji use influences students' impressions of credibility, immediacy, and liking. *Education and Information Technologies*, *28*(5), 6075-6092. https://doi.org/10.1007/s10639-022-11421-w

Varol, O., & Uluturk, I. (2019). Deception strategies and threats for online discussions. *ArXiv Preprint ArXiv:1906.11371*. https://doi.org/10.48550/arXiv.1906.11371

Vrontis, D., Makrides, A., Christofi, M., & Thrassou, A. (2021). Social media influencer marketing: A systematic review, integrative framework and future research agenda. *International Journal of Consumer Studies*, *45*(4), 617-644. https://doi.org/10.1111/ijcs.12647

Waddell, T. F., & Sundar, S. S. (2017). # thisshowsucks! The overpowering influence of negative social media comments on television viewers. *Journal of Broadcasting & Electronic Media*, *61*(2), 393-409. https://doi.org/10.1080/08838151.2017.1309414

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, *53*(2), 134-144. https://doi.org/10.1002/asi.10016

Weiner, Bernard (1974), Achievement Motivation and Attribution Theory. NewYork, NY: *General Learning Press.*

Wells, J. D., Valacich, J. S., & Hess, T. J. (2011). What signal are you sending? How website quality influences perceptions of product quality and purchase intentions. *MIS Quarterly*, 373-396. https://doi.org/10.2307/23044048

Wen, C., Tan, B. C., & Chang, K. T. T. (2009). Advertising effectiveness on social network sites: an investigation of tie strength, endorser expertise and product type on consumer purchase intention. *ICIS 2009 Proceedings*, 151. http://aisel.aisnet.org/icis2009/151

West, L. E. (2015). Responding (or not) on Facebook: A sociolinguistic study of liking, commenting, and other reactions to posts. *Georgetown University.*

Westerman, D., Spence, P. R., & Van Der Heide, B. (2014). Social media as information source: Recency of updates and credibility of information. *Journal of Computer-Mediated Communication, 19*(2), 171-183. https://doi.org/10.1111/jcc4.12041

Woollacott, E. (2019). Amazon's fake review problem is getting worse. Forbes. https://www.forbes.com/sites/emmawoollacott/2019/04/ 16/amazons-fakereview-problem-is-getting-worse/#de13d66195 f5. Accessed: 03.12.2021

Xiao, M., Wang, R., & Chan-Olmsted, S. (2018). Factors affecting YouTube influencer marketing credibility: a heuristic-systematic model. *Journal of Media Business Studies*, *15*(3), 188-213. https://doi.org/10.1080/16522354.2018.1501146

Xie, K. L., Zhang, Z., & Zhang, Z. (2014). The business value of online consumer reviews and management response to hotel performance. *International Journal of Hospitality Management*, *43*, 1-12. https://doi.org/10.1016/j.ijhm.2014.07.007

Xu, Q. (2014). Should I trust him? The effects of reviewer profile characteristics on eWOM credibility. *Computers in Human Behavior*, *33*, 136-144. https://doi.org/10.1016/j.chb.2014.01.027

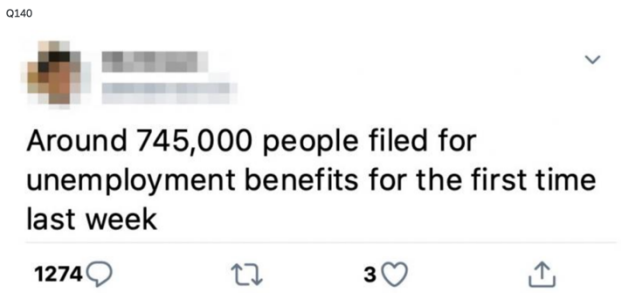
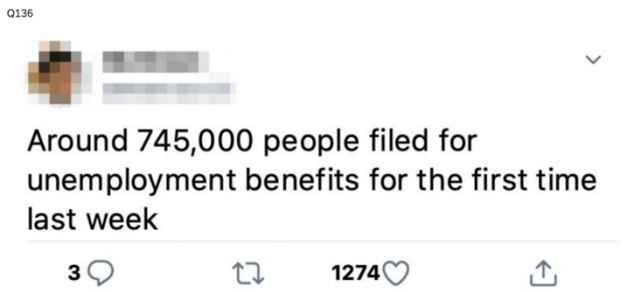
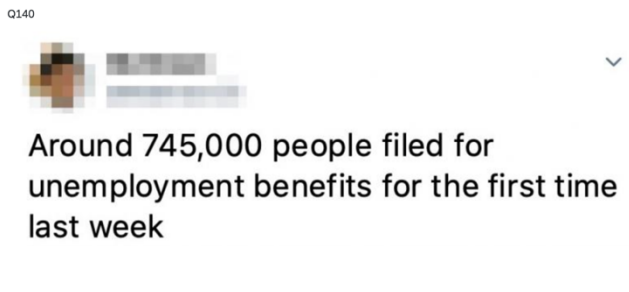
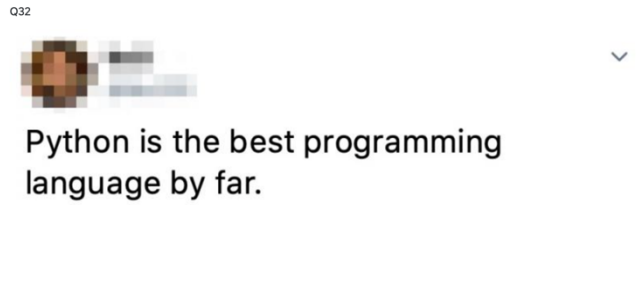
Zaichkowsky, J. L. (1994). The personal involvement inventory: Reduction, revision, and application to advertising. *Journal of Advertising*, *23*(4), 59-70. https://doi.org/10.1080/00913367.1943.10673459

Zhang, Z., Zhang, Z., & Li, H. (2015). Predictors of the authenticity of Internet health rumors. *Health Information & Libraries Journal*, 32(3), 195-205. https://doi.org/10.1111/hir.12115

Zhuang, M., Cui, G., & Peng, L. (2018). Manufactured opinions: The effect of manipulating online product reviews. *Journal of Business Research, 87*, 24–35. https://doi.org/10.1016/j.jbusres.2018. 02.016

WEB APPENDIX

Web Appendix 1: Sample stimuli (Study 1).

<table>
<tr><td colspan="2">High Ratio vs. Low Ratio Conditions</td></tr>
<tr><td> 
</td><td> 
</td></tr>
<tr><td> 
</td><td> 
</td></tr>
<tr><td colspan="2">Control Conditions</td></tr>
<tr><td> 
</td><td> 
</td></tr>
</table>

Web Appendix 2: Some of the stimuli used in Study 2.

	High ratio	Low ratio
Likes only	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Likes**: 3	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Likes**: 15
Comments only	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Comments**: 15	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Comments**: 3
Likes and comments	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Comments**: 15 **Likes**: 3	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Comments**: 3 **Likes**: 15
Ratio, likes, and comments	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Comments**: 15 **Likes**: 3 **Ratio (Comments:Likes)**: 5:1	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Comments**: 3 **Likes**: 15 **Ratio (Comments:Likes)**: 1:5
Ratio only	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Ratio (Comments:Likes)**: 5:1	**Company XY** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **Ratio (Comments:Likes)**: 1:5

Web Appendix 3: Study WA3 – Alternative explanations.

The objective of this study is to shed light on the process by which the ratio affects perceived credibility of a social media post in a controlled environment. This study was pre-registered at aspredicted.org (Web Appendix 9).

Procedure

We recruited 504 (M_{Age} = 39.5, SD_{Age} = 13.4, 42.2% male) participants from Mturk in exchange for a monetary payment that all participants who completed the study received. Participants of the previous study were excluded.

This study used a between-subjects design with five levels of the ratio. We manipulated the ratio by varying the number of likes of that tweet across conditions. The very low condition had a ratio of 0.011 (110 comments, 10,000 likes), the low condition had a ratio of 0.11 (110 comments, 1,000 likes), the medium condition had a ratio of 1.10 (110 comments, 100 likes), the high condition a ratio of 11.0 (110 comments, 10 likes), and the very high a ratio of 110 (110 comments, 1 likes). As we showed in Study 2, the effect is coming from the ratio itself and not the number of likes or comments; therefore, our manipulation of the number of likes, is basically a mere manipulation of ratio. It is worth mentioning that the content of the comments was not visible to the participants.

Each of the participants saw one single tweet. The content of the tweets was identical across conditions. We chose a post about the non-ionizing radiation and its effect on electro-hypersensitivity, such that prior knowledge, and thus knowledge about the credibility of the tweet's content, was unlikely to play a role. The manipulations are shown below (Figure 12).

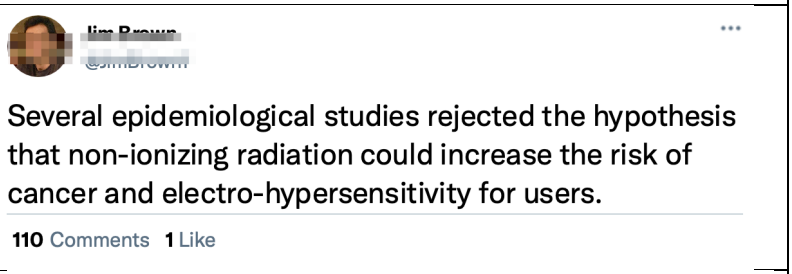
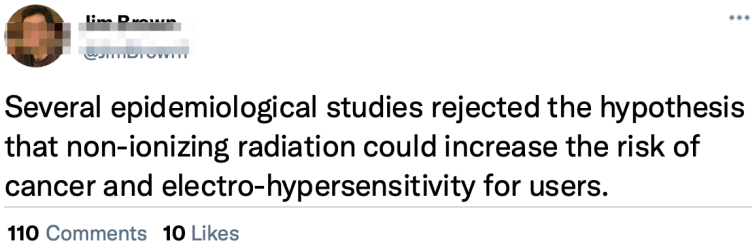
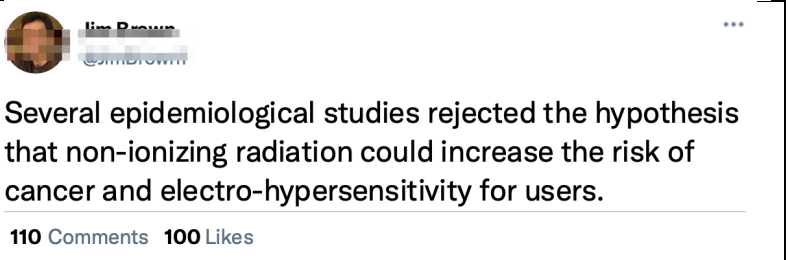
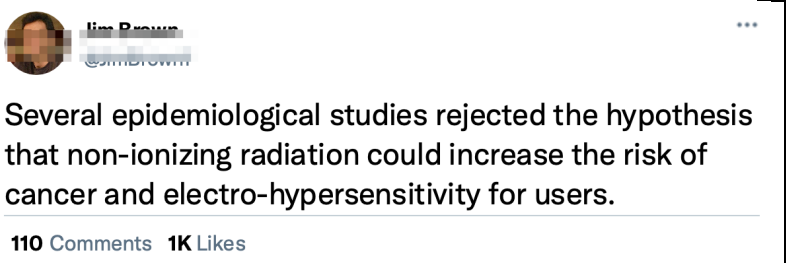
Ratio	*Stimulus*
110	
11	
1.1	
0.11	

Fig. 12. Very High vs. High vs. Medium vs. Low Ratio Conditions

Measurement

We measured perceived degree of consensus as explicit and implicit measures. To gauge explicit consensus, participants rated the percentage of the comments they believed to agree with the content of the tweet. Implicit consensus was measured using a single-item Inclusion of Other in the Self (IOS) scale (Aron et al. 1992). We also used overall perceived

consensus using a four-item scale, that reflects prior literature about perceived consensus (Cronbach a = .92).

We measured the perceived credibility with the same 9-item scale used in previous studies (Cronbach a = .96).

To gain further insight into the evaluation process, we also measured other probable mediators. The reason for this action is more exploratory than explanatory. Firstly, we measured positive and negative affect using the 20-item Positive and Negative Affect Scale, which ranges from 1 ('very slightly or not at all') to 5 ('extremely') (PANAS; Watson & Tellegen 1988). This was done to check whether the number of comments and likes could have any significant effect on the users' emotions, which could make our findings endogenous. The Cronbach alpha for positive affect was .92 and for negative affect was .92.

We also measured ambiguity using the 3-item Attitude Toward the Ad (Confusion) Scale (Ewing et al., 2005). This was done to check whether a high number of likes and comments could confuse users since ratioing does not occur frequently. The Cronbach alpha for this measure was .72.

Source credibility was measured using the modified version of the 9-item Spokesperson Credibility Scale (Eisend, 2006), as this construct is highly correlated with perceived credibility of the content. The Cronbach alpha for this measure was .96.

Involvement with the tweet was measured on a 10-item scale (Zaichkowsky, 1994). This was done to check whether exposure to something controversial with a high number of comments, which is again not frequent on social media platforms, can affect the way users want to be, or feel, involved in the topic. One of the justifications for measuring involvement is the probable reaction of the users due to the incongruency that they may find in the post because of the rare ratio of the post (Brehm, 1966). The Cronbach alpha for this measure was .96.

For the same reason, consumers' interest in the tweet was also measured using a modified version of the 3-item Interest in the Advertisement (Maclnnis et al., 2002). This was done as the heat created by the high number of likes and comments could make the topic of the post more interesting. The Cronbach alpha for this measure was .72. All measures were collected in randomized order.

All participants indicated their age (in groups), gender, and experience in using social media.

Results

Perceived credibility. A one-way ANOVA ($F(4, 499) = 9.72$, $p < .001$, $\eta^2 = .072$, $M_{0.011} = 4.71$, $SD = 1.22$, $M_{0.11} = 4.62$, $SD = .97$, $M_{1.1} = 4.65$, $SD = 1.08$, $M_{11} = 4.16$, $SD = 1.39$, $M_{110} = 3.83$, $SD = 1.48$) with conditions as independent variable and perceived credibility as dependent variable showed that conditions had a significant effect on perceived credibility. Thus, higher ratios significantly reduced the perceived credibility of the post. The results are shown in Figure 13.

Explicit perceived consensus. A one-way ANOVA ($F(4, 499) = 15.42$, $p < .001$, $\eta^2 = .11$, $M_{0.011} = 3.09$, $SD = 1.19$, $M_{0.11} = 3.00$, $SD = 1.04$, $M_{1.1} = 3.45$, $SD = 1.33$, $M_{11} = 2.16$, $SD = 1.10$, $M_{110} = 2.30$, $SD = 2.06$) with conditions as independent variable and explicit perceived consensus as dependent variable showed that conditions had a significant effect on perceived credibility. Thus, higher ratios significantly reduced the explicit consensus of the comments with the content of the post.

Perceived degree of consensus. A one-way ANOVA ($F(4, 499) = 25.05$, $p < .001$, $\eta^2 = .17$, $M_{0.011} = 4.58$, $SD = 1.28$, $M_{0.11} = 4.51$, $SD = 1.20$, $M_{1.1} = 4.68$, $SD = 1.32$, $M_{11} = 3.34$, $SD = 1.39$, $M_{110} = 3.31$, $SD = 1.43$) with conditions as independent variable and implicit perceived consensus as dependent variable showed that conditions had a significant effect on

perceived credibility. Thus, higher ratios significantly reduced the implicit consensus of the comments with the content of the post.

Implicit perceived consensus. A one-way ANOVA ($F(4, 499) = 26.86$, $p < .001$, $\eta^2 = .18$, $M_{0.011} = 3.97$, $SD = 1.47$, $M_{0.11} = 4.00$, $SD = 1.23$, $M_{1.1} = 4.30$, $SD = 1.50$, $M_{11} = 2.90$, $SD = 1.39$, $M_{110} = 2.65$, $SD = 1.49$) with conditions as independent variable and the degree of perceived consensus as dependent variable showed that conditions had a significant effect on perceived credibility. Thus, higher ratios significantly reduced the degree of consensus of the comments with the content of the post.

The results of the one-way ANOVAs for all other alternative explanations are as follows.

PANAS positive. The one-way ANOVA was non-significant ($F(4, 499) = 1.72$, $p = .14$, $\eta^2 = .01$, $M_{0.011} = 2.11$, $SD = .90$, $M_{0.11} = 1.86$, $SD = .78$, $M_{1.1} = 1.97$, $SD = .79$, $M_{11} = 1.91$, $SD = .82$, $M_{110} = 1.85$, $SD = .88$.

PANAS negative. The one-way ANOVA was non-significant ($F(4, 499) = .81$, $p = .52$, $\eta^2 = .01$, $M_{0.011} = 1.29$, $SD = .59$, $M_{0.11} = 1.18$, $SD = .35$, $M_{1.1} = 1.23$, $SD = .50$, $M_{11} = 1.26$, $SD = .50$, $M_{110} = 1.22$, $SD = .42$.

Interest in the social media post. The one-way ANOVA was significant ($F(4, 499) = 4.97$, $p = .001$, $\eta^2 = .04$, $M_{0.011} = 1.74$, $SD = 1.28$, $M_{0.11} = 1.38$, $SD = 1.23$, $M_{1.1} = 1.67$, $SD = 1.09$, $M_{11} = 1.24$, $SD = 1.25$, $M_{110} = 1.12$, $SD = 1.25$.

Involvement with the social media post. The one-way ANOVA was significant ($F(4, 499) = 7.98$, $p < .001$, $\eta^2 = .06$, $M_{0.011} = 4.44$, $SD = 1.25$, $M_{0.11} = 4.23$, $SD = 1.14$, $M_{1.1} = 4.40$, $SD = 1.06$, $M_{11} = 3.67$, $SD = 1.34$, $M_{110} = 3.83$, $SD = 1.31$.

Perceived source credibility. The one-way ANOVA was significant ($F(4, 499) = 7.64$, $p < .001$, $\eta^2 = .06$, $M_{0.011} = 4.85$, $SD = 1.14$, $M_{0.11} = 4.64$, $SD = 1.02$, $M_{1.1} = 4.75$, $SD = 1.05$, $M_{11} = 4.21$, $SD = 1.28$, $M_{110} = 4.14$, $SD = 1.34$.

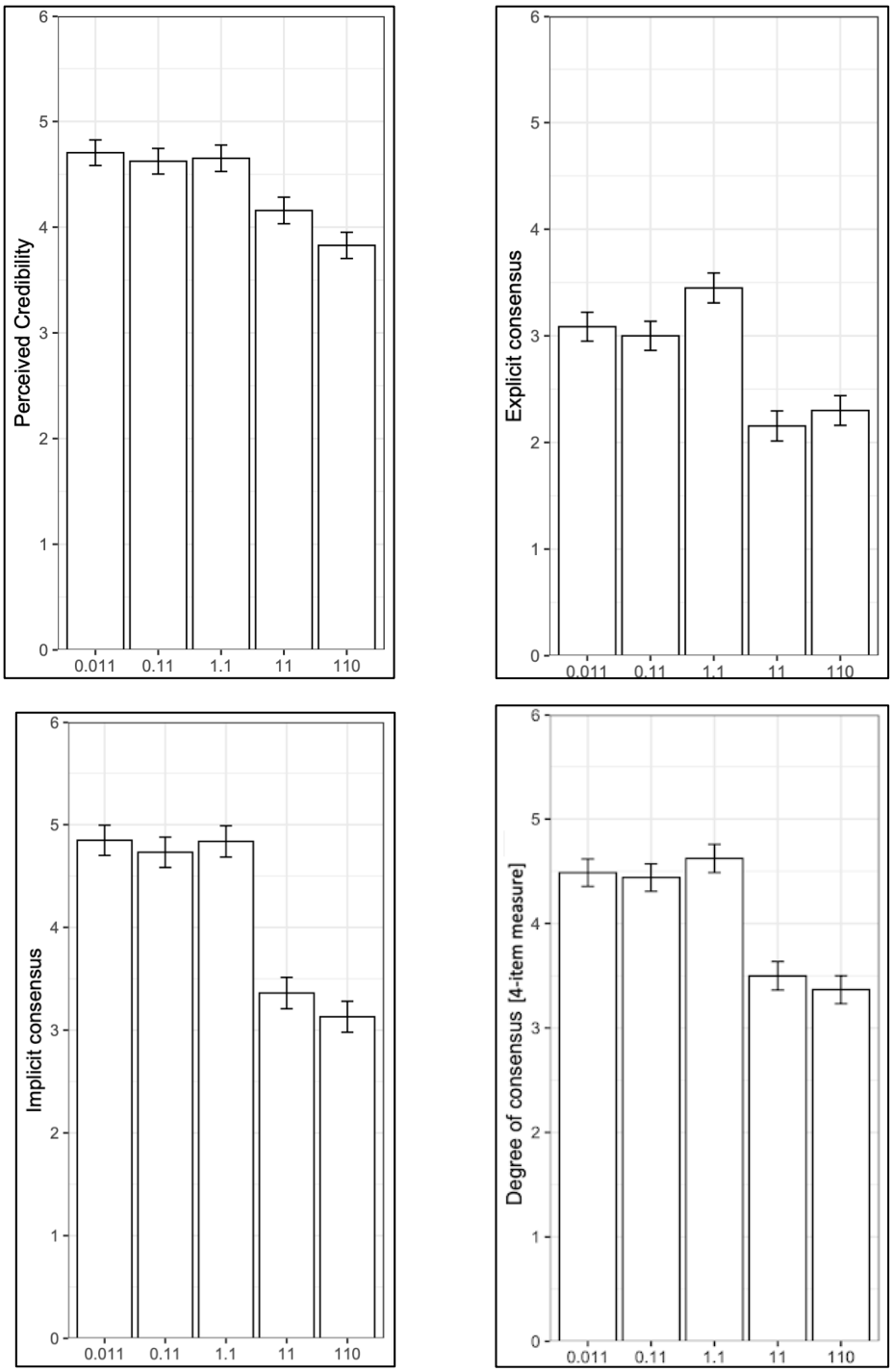


Fig. 13. A High Ratio Reduces the (A) Perceived Credibility, (B) Explicit Consensus, (C) Implicit Consensus, and (D) Perceived Degree of Consensus.

Mediation by perceived degree of consensus. We ran several mediation models to test whether the degree of consensus not only varied across conditions but also whether it predicted the perceived credibility of a social media post. First, we ran individual mediations with each of the three measures of consensus entered individually as mediators.

Using the PROCESS Model (Hayes, 2012; Model 4, $N_{Bootstraps}$ = 10,000) with conditions as the independent variable, implicit consensus as the mediator, and perceived credibility as the dependent variable, revealed a significant indirect effect (B = -.14, SE = .02, CI95% = [-.19; -.10]), while the direct effect remained significant (B = -.08, SE = .04, CI95% = [-.15; -.004]). Using the same model, but with the four-item measure of the perceived degree of consensus again yielded a significant indirect effect (B = -.18, SE = .02, CI95% = [-.23; -.13]) while the direct effect of the ratio on perceived credibility became non-significant (B = -.04, SE = .03, CI95% = [-.11; .03]). Finally, using the same model, but with the explicit measure of perceived degree of consensus again yielded a significant indirect effect (B = -.09, SE = .02, CI95% = [-.14; -.05]), while the direct effect again remained significant (B = -.13, SE = .04, CI95% = [-.20; -.06]). Overall, the mediations showed that, independent of the measure of perceived degree of consensus, this variable negatively mediated the effect of the ratio on perceived credibility of the post.

Alternative explanations. We conducted another mediation model (Model 4, $N_{Bootstraps}$ = 10,000) that included all potential alternative explanations as additional mediators in addition to the four-item measure of the perceived degree of consensus. The indirect effect via degree of consensus remained significant (B = -.06, SE = .01, CI95% = [-.09; -.04]) while at the same time, the indirect effect via perceived source credibility was significant (B = -.14, SE = .03, CI95% = [-.19; -.08]). The direct effect became non-significant as well (B = -.02, SE = .02, CI95% = [-.06; .03]). All other indirect effects were non-significant, even with using other measurements of perceived consensus (see Tables 4, 5, and 6 in Web Appendix 4). While source credibility arguably is closely related to the overall credibility (and thus was likely to turn out as a significant mediator) none of the other constructs mediated the effect of ratio on perceived credibility. This provides strong evidence for the degree of consensus as an explanation for the lower perceived credibility of the social media post.

Discussion

Study WA3 provides evidence that the perceived degree of consensus caused a high ratio to be less credible than a low ratio. The study also provides evidence that when the comment-to-like ratio exceeds a certain level, perceived credibility is reduced significantly, while up to an equal number of comments and likes, the perceived credibility does not decrease. This study also rules out several alternative explanations, i.e., sentiment of the post, involvement with the post, and ambiguity.

Web Appendix 4: Table of indirect effects (from Study WA3)

Table 4

Indirect effects for the four-item measure of perceived consensus

		B	**SE**	**95% Lower**	**95% Upper**
Total		-.20	.04	-.27	-.13
Indirect	Four-item consensus	-.06	.01	-.09	-.03
	interest	.01	.01	-.01	.02
	involvement	-.01	.01	-.04	.003
	PANAS positive	.001	.003	-.004	.01
	PANAS negative	-.0001	.001	-.003	.002
	User credibility	-.13	.03	-.19	-.08
Direct		-.02	.02	-.06	.03

Table 5

Indirect effects for the measure of explicit consensus.

		B	**SE**	**95% Lower**	**95% Upper**
Total		-.18	.03	-.25	-.11
Indirect	Explicit consensus	-.03	.01	-.05	-.01
	interest	.004	.005	-.01	.02
	involvement	-.02	.01	-.04	.001
	PANAS positive	.001	.002	-.004	.01
	PANAS negative	-.0001	.001	-.002	.002
	User credibility	-.14	.03	-.20	-.08
Direct		-.04	.02	-.09	.003

Table 6

Indirect effects for the measure of implicit consensus.

		B	**SE**	**95% Lower**	**95% Upper**
Total		-.19	.03	-.26	-.13
Indirect	Implicit consensus	-.04	.01	-.07	-.02
	interest	.004	.01	-.01	.02
	involvement	-.02	.01	-.04	.001
	PANAS positive	.001	.003	-.004	.01
	PANAS negative	-.0001	.001	-.003	.002
	User credibility	-.14	.03	-.20	-.08
Direct		-.03	.02	-.07	.02

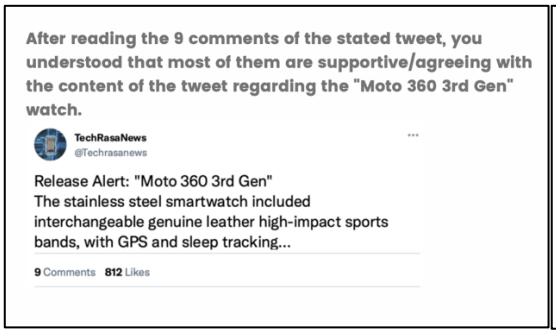
Web Appendix 5: Stimuli used in Study 3.

Manipulation of the ratio

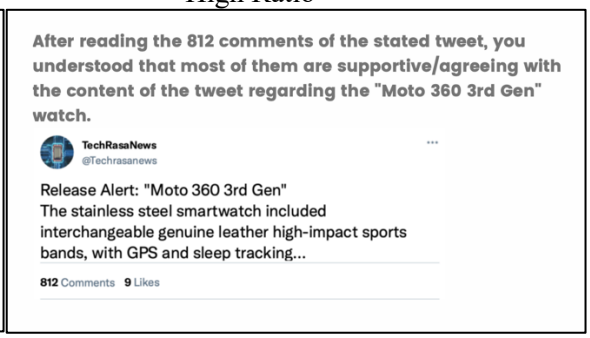
Low ratio	High ratio
TechRasaNews @Techrasanews Release Alert: "Withings Steel HR Sport" Multisport hybrid smartwatch, with connected GPS Easy activity and sleep tracking, with interchangeal bands... **13** Comments **19** Likes	TechRasaNews @Techrasanews Release Alert: "Withings Steel HR Sport" Multisport hybrid smartwatch, with connected GPS Easy activity and sleep tracking, with interchangeal bands... **13** Comments **19** Likes
TechRasaNews @Techrasanews Release Alert: "Moto 360 3rd Gen" The stainless steel smartwatch included interchangeable genuine leather high-impact sports bands, with GPS and sleep tracking... **9** Comments **812** Likes	TechRasaNews @Techrasanews Release Alert: "Moto 360 3rd Gen" The stainless steel smartwatch included interchangeable genuine leather high-impact sports bands, with GPS and sleep tracking... **812** Comments **9** Likes
TechRasaNews @Techrasanews Release Alert: "Booland Minimalist Smart Watch" Hybrid Smartwatch Fitness Tracker for Men & Wom with activity/sleep tracking and notification center.. **10** Comments **15** Likes	TechRasaNews @Techrasanews Release Alert: "Booland Minimalist Smart Watch" Hybrid Smartwatch Fitness Tracker for Men & Wom with activity/sleep tracking and notification center. **10** Comments **15** Likes

Manipulatio for “Present Consensus Cue”

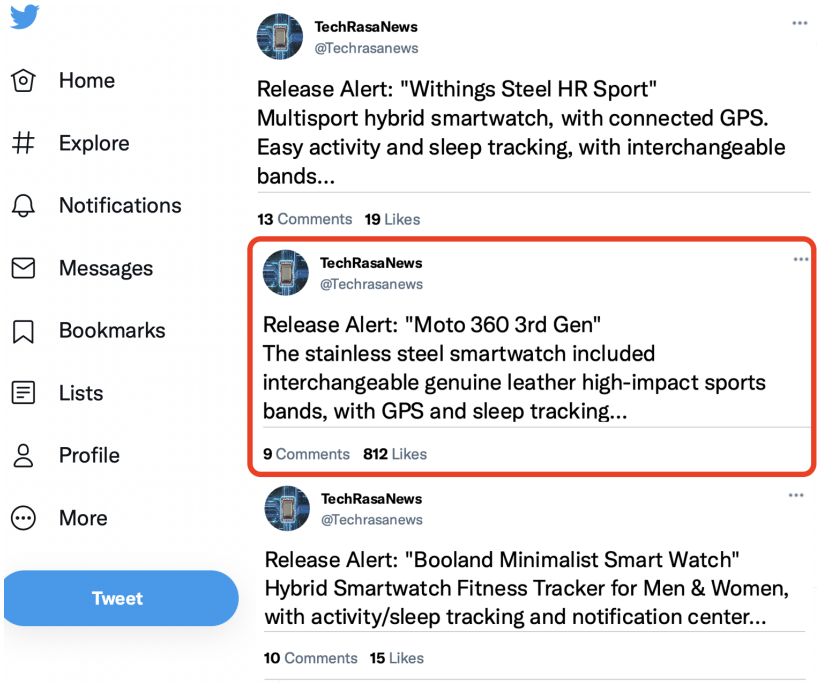
Low Ratio



High Ratio



Web Appendix 6: Example of the stimuli used in Study 4.



Web Appendix 7: Sample stimuli (Study 5).

High ratio × feedback solicited	**Company Alpha** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. Please share your feedback about our smart bottle in the comment section. **#Comments=15** **#Likes= 3**
High ratio × Feedback not-solicited	**Company Alpha** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **#Comments=15** **#Likes= 3**
Normal ratio × feedback not-solicited	**Company Alpha** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. **#Comments = 3** **#Likes = 15**
Normal ratio × Feedback solicited	**Company Alpha** We are glad to introduce our new product. The smart bottle is a quality product and was launched in February 2022. It can measure your water intake over the whole day. Our customers are highly satisfied with this new bottle. Please share your feedback about our smart bottle in the comment section. **#Comments= 3** **#Likes= 15**

Web Appendix 8: Study WA8 - Effects of the first comment.

Popular social media platforms like Twitter and Instagram display a single comment, typically the most relevant one to the user, below each post on the timeline. In this study, we aim to examine whether the presentation of a positive or negative first comment can overshadow the ratioing cue, as it may be perceived as a simpler cue compared to the ratio of two distinct indicators. This study was also pre-registered at aspredicted.org (Web Appendix 9)

Procedure

We recruited 603 (M_{Age} = 42.1, SD_{Age} = 13.8, 47.6% male) participants from MTurk in exchange for a monetary payment that all participants who completed the study received. Participants of the previous studies were also excluded.

This study employed a between-subjects design of 2 × 3 (high vs. normal ratio) × (negative vs. positive vs. neutral first comment). The content, social media platform, and ratio manipulations in this study were identical to those in Study 5. However, we added a comment under the post to implement the three comment manipulations. As shown below, we used "Is there also an IOS app for it?" for the 'neutral' condition, "I actually did not like it" for the 'negative' condition, and "I actually liked it" for the 'positive' condition as the first comment displayed beneath the post.

Results

Manipulation Check. A two-way ANOVA with comment and ratio as independent variables and perceived valence of the comment as dependent variable, showed a significant effect between our manipulation and the valence perception ($F(2,597) = 448.21$, $p < 001$, $\eta^2 = .60$). This result shows that our manipulation in this study worked. It is worth mentioning that for measuring the perceived valence of the comments, we used a one-item Likert-type scale,

asking the participants to indicate to what extent they believe that the comment under the post was supportive toward it. (from 1= “Totally disagree” to 5= “Totally agree”).

Perceived credibility. A two-way ANOVA with comment and ratio as independent variables and perceived credibility as dependent variable ($M_{HighRatio\_NegativeComment}$ = 4.20, SD = 1.29, $M_{HighRatio\_NormalComment}$ = 4.73, SD = 1.09, $M_{HighRatio\_PositiveComment}$ = 4.83, SD = 1.07 $M_{NormalRatio\_NegativeComment}$ = 4.62, SD = .92, $M_{NormalRatio\_NormalComment}$ = 5.11, SD = .87, $M_{NormalRatio\_PositiveComment}$ = 5.21, SD = .82) showed that the presence of the first comment on the post has no significant moderation effect on the effect of ratio on perceived credibility ($F(2, 597) = .02$, $p = .98$, $\eta^2 = <.001$). Ratio ($F(1, 597) = 21.49$, $p < .001$, $\eta^2 = .03$) and presence of the first comment had a significant main effect ($F(2, 597) = 24.13$, $p < .001$, $\eta^2 = .06$). The results are shown in Figure 14.

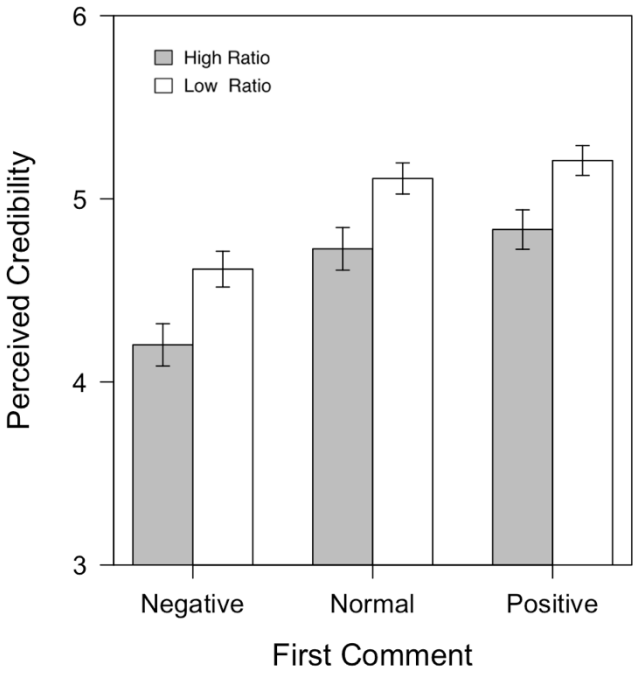


Fig. 14. The Ratioing Effect on Perceived Credibility Is Robust Against the First Comment of the Post.

Discussion

This study demonstrates that our model remains robust against the first comment, whether it is positive, negative, or neutral. The comment did not have any boundary effect on the proposed model regarding ratioing. This study once again underscores that despite the availability of simpler cues, people are inclined to rely on ratioing as a more robust (albeit more complex) cue.

Web Appendix 9: Links of all pre-registered studies.

Study	Pre-Registration Link
Study 2	https://aspredicted.org/3BC_RRG
Study Web Appendix 3	https://aspredicted.org/BGY_78W
Study 3	https://aspredicted.org/F75_WYG
Study 5	https://aspredicted.org/NZD_K95
Study Web Appendix 8	https://aspredicted.org/45Z_Q22

Web Appendix 10. Analysis of all five conditions (Study 2).

A two-way ANOVA using the ratio and the formats as independent variables and degree of consensus as dependent variable was significant ($F_{Interaction}(4, 992) = 9.86$, $p < .001$, $\eta2 = .04$, $F_{Format}(4, 992) = 5.46$, $p < .001$, $\eta2 = .02$, $F_{Ratio}(1, 992) = 43.98$, $p < .001$, $\eta2 = .04$, $M_{NumberOfComments\_High} = 4.14$, SD = 1.19, $M_{NumberOfLikesAndComments\_High} = 3.14$, SD = 1.43, $M_{NumberOfCommentsLikesAndRatio\_High} = 2.71$, SD = 1.23, $M_{NumberOfLikes\_High} = 3.79$, SD = 1.74, $M_{Ratio\_High} = 3.38$, SD = 1.55, $M_{NumberOfComments\_Low} = 3.95$, SD = 1.48, $M_{NumberOfLikesAndComments\_Low} = 4.11$, SD = 1.38, $M_{NumberOfCommentsLikesAndRatio\_Low} = 4.21$, SD = 1.52, $M_{NumberOfLikes\_Low} = 4.15$, SD = 1.61, $M_{Ratio\_Low} = 3.76$, SD = 1.55). Planned contrasts showed that the effect of the high ratio was significant in the 'number of likes and comments' ($t = -4.72$, $p < .001$) and 'number of comments, likes, and ratio' ($t = -7.36$, $p < .001$) conditions, marginally significant for 'ratio' ($t = -1.78$, $p = .07$) and 'number of likes' ($t = -1.69$, $p = .09$)) conditions, while it was insignificant for the 'number of comments' ($t = .94$, $p = .35$) condition.

Another two-way ANOVA using the ratio and the formats as independent variables and perceived credibility as dependent variable was significant ($F_{Interaction}(4, 992) = 5.63$, $p < .001$, $\eta^2 = .02$, $F_{Format}(4, 992) = 4.35$, $p = .002$, $\eta^2 = .02$, $F_{Ratio}(1, 992) = 24.82$, $p < .001$, $\eta^2 = .02$, $M_{NumberOfComments\_High} = 4.90$, SD = .92, $M_{NumberOfLikesAndComments\_High} = 4.36$, SD = 1.08, $M_{NumberOfCommentsLikesAndRatio\_High} = 4.35$, SD = 1.17, $M_{NumberOfLikes\_High} = 5.05$, SD = 1.10, $M_{Ratio\_High} = 4.45$, SD = 1.27, $M_{NumberOfComments\_Low} = 4.96$, SD = .96, $M_{NumberOfLikesAndComments\_Low} = 4.86$, SD = 1.05, $M_{NumberOfCommentsLikesAndRatio\_Low} = 5.10$, SD = 1.00, $M_{NumberOfLikes\_Low} = 4.91$, SD = 1.10, $M_{Ratio\_Low} = 4.89$, SD = .89). Planned contrasts showed that the effect of the high ratio was significant in the 'number of likes and comments' ($t = -3.40$, $p < .001$), 'number of comments, likes, and ratio' ($t = -5.14$, $p < .001$), and 'ratio' ($t = -2.88$, $p = .004$) conditions, while it was insignificant for the 'number of comments' ($t = -$

.49, p = .65) and ‘number of likes’ (t =.92, p = .36) conditions. These results show that having either the number of likes or the number of comments may not affect the perceived credibility, while having the ratio (explicitly or implicitly) can affect the perceived credibility of the post through the mediation channel of perceived degree of consensus.

Web Appendix 11. Descriptive Statistics of the Studies.

Study	***Construct***	***Min***	***Max***	***Mean***	***SE***
Field Study	*Number of likes*	0	151700	1248.04	197.41
	Number of comments	1	19200	55.02	14.64
	Ratio	.001	96	.25	.07
	Degree of consensus	1	9	6.39	.05
Study 1	*Perceived credibility*	1	7	4.49	1.59
Study 2	*Perceived credibility*	1.4	7	4.79	.03
	Perceived consensus	1	7	3.74	.05
Study 3	*Perceived credibility*	1	7	5.22	.05
	Perceived consensus	1	7	4.31	.08
	Purchase intention	1	7	4.07	.08
Study 4	*Perceived credibility*	1	7	5.36	.02
	Choice Intention	0	1	.51	.02
Study 5	*Perceived credibility*	1	7	4.75	.06
	Degree of consensus	1	7	3.78	.07

Web Appendix 12. Decision Tree Analysis

To unravel the hierarchy of importance among the various potential predictors of a post's perceived credibility, we conducted a comprehensive decision tree analysis utilizing our field data. Our analysis encompassed all coded parameters of the tweets, encompassing factors such as the number of likes, comments, retweets, URLs, hashtags, media, as well as the application of linguistic analysis through the LIWC, PARA, and Harvard dictionaries. Furthermore, we fine-tuned the Complexity Parameter (CP) by evaluating the Root Mean Square Error (RSME) across different CP values, specifically .001, .01, .1, .2, .5, and 1. This approach allowed us to systematically assess the impact of various predictors on perceived credibility. The decision tree with depth of 12 levels can be found in Figure 15.

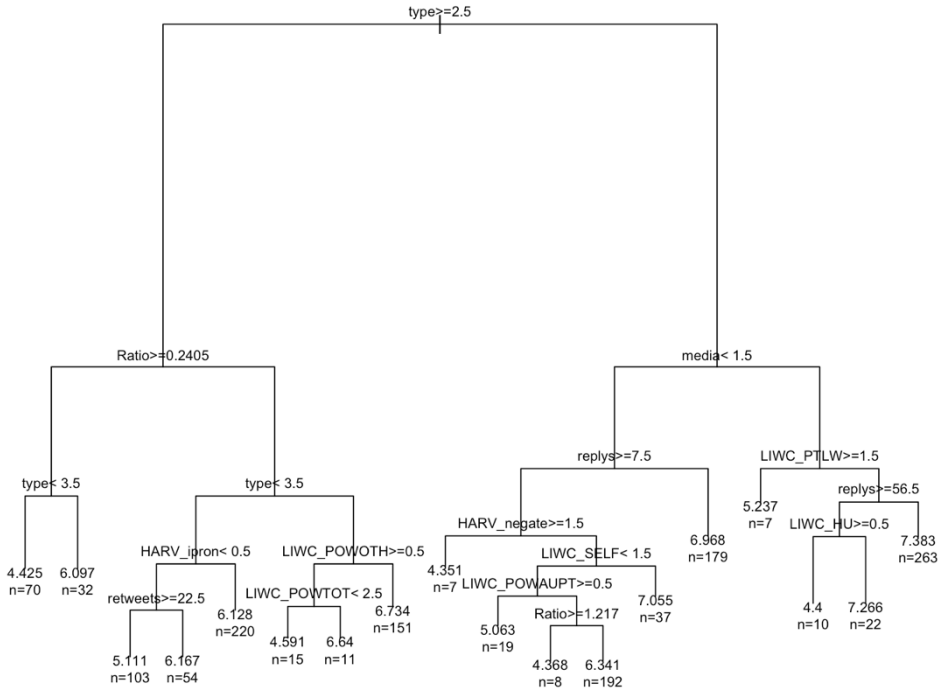


Fig. 15. Hierarchical Predictors of Level of Agreement (Depth Limited to 12 Levels)

Web Appendix 13: Links of all pre-registered studies.

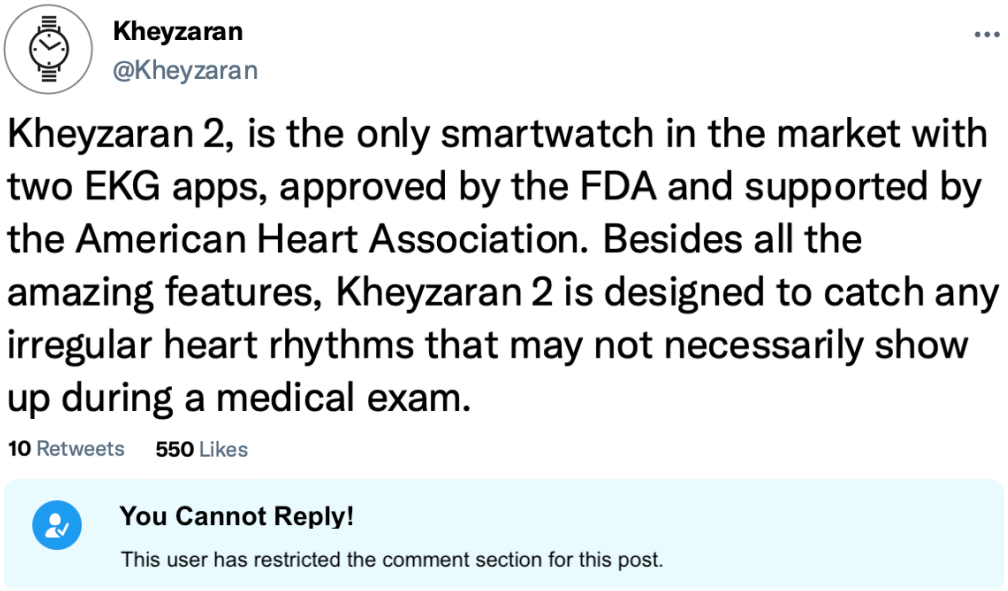
Study	Pre-Registration Link
Study 1	https://aspredicted.org/T7J_85R
Study 2	https://aspredicted.org/R9W_MLD
Study 3	https://aspredicted.org/MYV_QRH
Study 4	https://aspredicted.org/FB4_2DJ
Study 5	https://aspredicted.org/2HF_7QQ
Study 6	https://aspredicted.org/C8M_QGH
Study WA1	https://aspredicted.org/MXV_HKB

Web Appendix 14. Descriptive Statistics of the Studies.

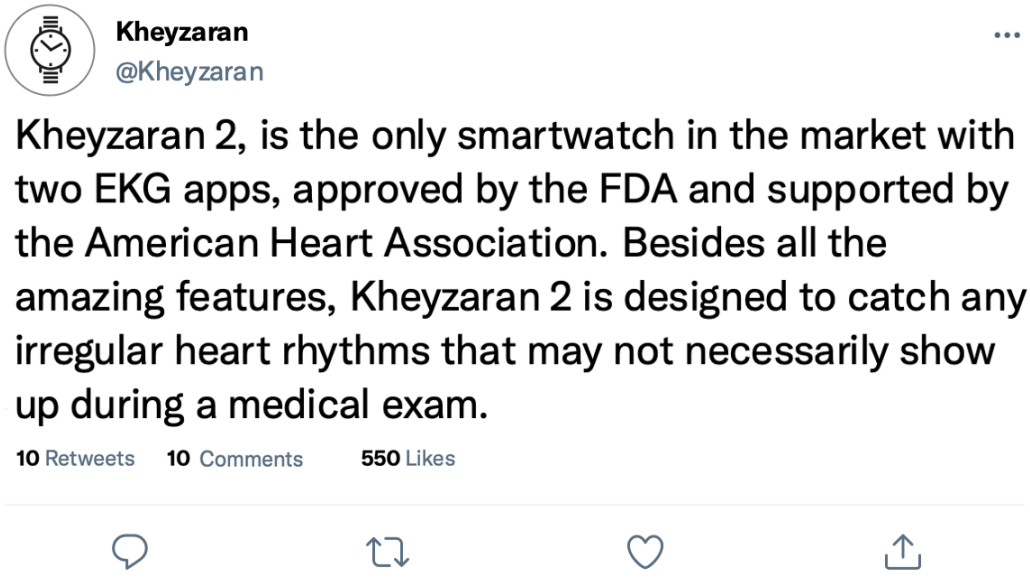
Study	***Construct***	***Min***	***Max***	***Mean***	***SE***
Field Study	*Attitude (RA1)*	0	10	6.26	.10
	Attitude (RA2)	0	10	5.99	.09
	Likes	0	53880	735.20	203.70
	Comments	0	3089	38.93	12.45
	Retweets	0	11660	162.20	42.67
	Quote-Retweets	1	2459	17.57	6.67
Study 1	*Perceived Censorship*	1	7	3.24	.12
	Suspicion	1	6.89	3.97	.08
	Attitude Toward the brand	-.4	5.4	2.83	.08
Study 2	*Perceived Censorship*	1	7	3.71	.13
	Suspicion	1	7	4.79	.09
	Attitude Toward the brand	-.6	5.2	2.03	.10
	Consistency	1	7	4.62	.10
	Customer-responsiveness	1	7	3.93	.10
	Perceived cyberbullying	1	6.2	3.25	.09
Study 3	*Perceived Censorship*	1	7	4.39	.09
	Suspicion	1	7	4.97	.06
	Attitude Toward the brand	-.6	5.4	1.75	.07
Study 4	*Perceived Censorship*	1	7	4.59	.08
	Suspicion	1	7	4.71	.06
	Attitude Toward the brand	-.6	5.4	2.50	.06
Study 5	*Perceived Censorship*	1	7	3.98	.09
	Suspicion	1	7	4.72	.07
	Attitude Toward the brand	-.6	5.4	1.88	.07
Study 6	*Perceived Censorship*	1	7	4.21	.08
	Suspicion	1	7	4.44	.05
	Attitude Toward the brand	-.6	5.4	2.29	.05
Study WA1	*Perceived Censorship*	1	7	3.40	.09
	Suspicion	1	7	4.17	.06
	Attitude Toward the brand	-.6	5.4	2.56	.06

Web Appendix 15. Stimuli Used in the Study 1

Restricted Condition:

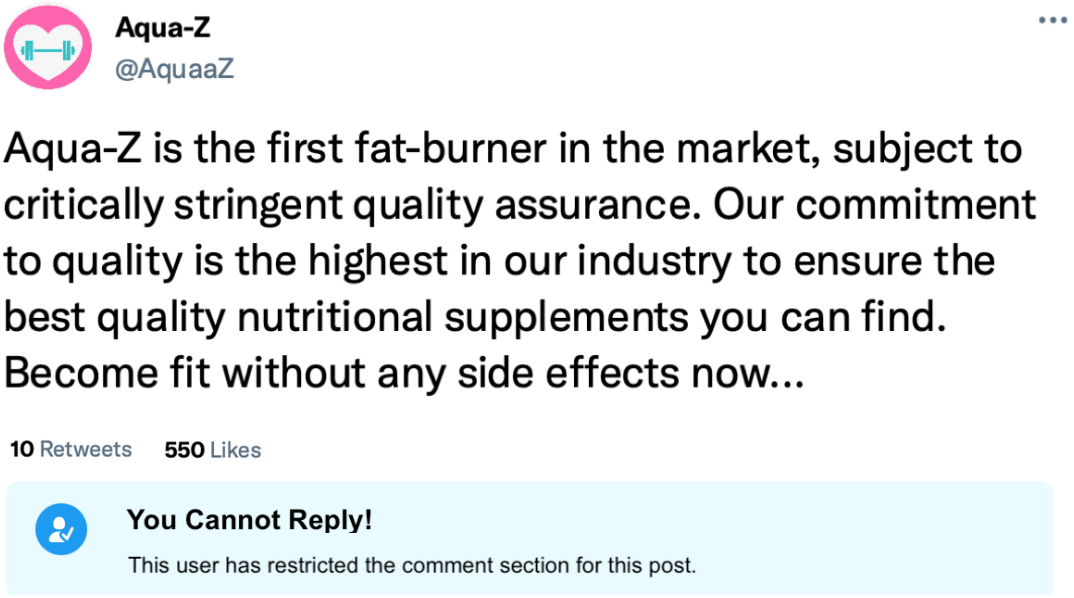


Control Condition:



Web Appendix 16. Stimuli Used in the Study 2

Restricted Condition:



Control Condition:

Web Appendix 17. Stimuli Used in the Study 4

Restricted x High Within-Consistency Conditions

Control x Low Within-Consistency Conditions

Web Appendix 18. Analysis of the Pre-Study

In a preliminary exploratory study conducted on MTurk (N=100, M_{Age}=36.3, SD_{Age}=10.8, 59% male), individuals were exposed to a social media post (a tweet) featuring a new product. Notably, the tweet had restricted commenting capabilities. Subsequently, participants were prompted to provide their insights into the reasoning behind the brand's decision to restrict comments through an open-ended question.

Participants were then queried about whether they had initially noticed the comment restriction when they first viewed the tweet. Following this, we asked participants to elaborate on their emotions and perceptions toward the brand that imposed these comment restrictions. Lastly, we prompted participants to envision a scenario in which limiting comments might positively influence their perception of the brand. Additionally, we collected demographic information, including age, gender, and participants' level of experience in using Twitter.

For coding the responses to the open-ended questions, we used the help of a TA, blinded to the hypothesis.

Results

Firstly, the attention-check question assessed whether participants noticed the visual cue indicating that commenting was restricted on the tweet. The results showed that 70% of participants became aware of the comment restriction upon viewing the tweet, suggesting that the visual cue of this act is quite visible and strong, at least on Twitter.

Furthermore, the analysis of coded responses reveals that a majority of participants believe that the primary reasons for limiting comments are either trolling and hate speech or censorship. Specifically, 35% of responses cited trolling and hate speech, while 33% mentioned censorship. As for the question concerning perceptions and attitude shifts toward the user who restricted comments, 28% indicated no change in their feelings toward the post, 23% reported growing suspicious, 19% became disinterested, and 10% felt censored.

Intriguingly, when asked about identifying a positive aspect of comment limitations, 47% could not think of a beneficial scenario, while only 8% cited privacy concerns and 7% mentioned defense against trolls and hate speech.

This study suggests that although users recognize various reasons for comment restrictions—such as information privacy, trolling, and hate speech—they generally harbor negative feelings toward such actions.

Web Appendix 19. Study WA1

This study shifts its focus to exploring another potential moderator, which may align with our theoretical model. Specifically, this study examines whether the importance of social media posts could moderate the relationship between restricting the ability to comment on social media posts on consumers' suspicion and attitude toward the brand.

Essentially, we aim to determine whether there is an acceptable scenario in which restricting the comments for posts of lesser importance does not result in negative consequences.

Procedure

We recruited 400 participants (M_{Age} = 40.6, SD_{Age} = 12.4, 43.0% male) from Amazon Mechanical Turk ("MTurk") in exchange for monetary compensation. Moreover, individuals who had taken part in any of the preceding studies were deemed ineligible for participation in this study. Employing a 2×2 between-subjects design, the study encompassed two independent variables: Status (Restricted vs. Control) and the level of importance of the topic (Low vs. High).

The study's design closely resembles the previous study, revolving around the promotional content of the "Aqua-Z" brand. However, participants were randomly allocated to either the 'High' or 'Low' importance group. In the low-importance group, participants came across a tweet from Aqua-Z introducing their new smartwatch, emphasizing the product's aesthetic features. Conversely, in the high-importance group, the tweet from Aqua-Z highlighted the health monitoring features of the same smartwatch which may save the life of users. Importantly, it is noteworthy that the posts were randomly distributed across both the 'Restricted' and 'Control' conditions regarding the comment section status.

Measures

All measurements utilized in this study remain aligned with those employed in Study 5. These encompass attitude toward the brand (Cronbach's $\alpha = .91$), suspicion (Cronbach's $\alpha = .93$), perceived censorship, age, gender, level of experience in using social media, and a manipulation check tied to the restriction of the comments. Additionally, an extra manipulation check was introduced to gauge the effectiveness of the manipulation concerning the level of importance of the tweet. Participants were prompted to indicate their agreement with the statement: "*The user's post (Aqua-Z) held significant importance for me,*" using a scale ranging from totally disagree (0) to totally agree (7). All descriptive statistics for this study can be found in Appendix 14.

Results

Manipulation checks

A two-way ANOVA with the status manipulation (Restricted vs. Control) and level of importance (Low vs. High) as the independent variables, and the importance manipulation check as the dependent variable ($M_{Restricted\_HighImportance} = 2.83$, $SD = 1.51$, $M_{Control\_HighImportance} = 3.34$, $SD = 1.50$, $M_{Restricted\_LowImportance} = 2.45$, $SD = 1.34$, $M_{Control\_LowImportance} = 2.28$, $SD = 1.30$), reveals the effectiveness of the manipulation on the level of importance of the topic in this study, evident through the significant influence of the level of importance on its associated manipulation check ($F(1, 396) = 25.65$, $p < .001$, $\eta^2 = .06$).

Another two-way ANOVA was executed, using our status manipulation (Restricted vs. Control) and level of importance (Low vs. High) as the independent variables, while utilizing the status manipulation check as the dependent variable ($M_{Restricted\_HighImportance} = 3.75$, $SD = .78$, $M_{Control\_HighImportance} = 2.87$, $SD = 1.37$, $M_{Restricted\_LowImportance} = 3.65$, $SD = .87$, $M_{Control\_LowImportance} = 2.60$, $SD = 1.39$). The results showed that the status manipulation also worked as intended ($F(1, 396) = 72.78$, $p < .001$, $\eta^2 = .15$).

Moderation of level of importance

Conducting a two-way ANOVA, where the status manipulation and level of importance served as the independent variables, and perceived censorship as the dependent variable ($M_{Restricted\_HighImportance}$ = 4.08, SD = 1.75, $M_{Control\_HighImportance}$ = 2.36, SD = 1.25, $M_{Restricted\_LowImportance}$ = 4.42, SD = 1.84, $M_{Control\_LowImportance}$ = 2.70, SD = 1.43) shows no significant moderation effect of the level of importance on perceived censorship ($F(1, 396) = .00$, $p = .99$, $\eta^2 < .001$). However, the main effect of the post's status remains significant ($F(1, 396) = 116.85$, $p < .001$, $\eta^2 = .23$). In the high-importance conditions, restricting the comments significantly increased perceived censorship ($t = 7.66$, $p < .001$), and this effect remained significant in the low importance conditions ($t = 7.66$, $p < .001$) as well.

Another two-way ANOVA with our status manipulation and level of importance as the independent variables, and suspicion as dependent variable ($M_{Restricted\_HighImportance}$ = 3.99, SD = 1.06, $M_{Control\_HighImportance}$ = 3.52, SD = 1.00, $M_{Restricted\_LowImportance}$ = 4.67, SD = 1.09, $M_{Control\_LowImportance}$ = 4.49, SD = 1.16), again reveals no significant moderation effect of level of importance ($F(1, 396) = 1.74$, $p = .19$, $\eta^2 = .004$) on suspicion, while the main effect of post's status remains significant ($F(1, 396) = 8.79$, $p = .003$, $\eta^2 = .02$). In the high importance condition, limiting the comments had a significant effect on suspicion ($t = 3.08$, $p = .002$), while in the low important condition this effect was non-significant ($t = 1.22$, $p = .22$).

In the final two-way ANOVA, where the status manipulation and level of importance served as the independent variables, and attitude toward the brand as the dependent variable ($M_{Restricted\_HighImportance}$ = 2.63, SD = 1.07, $M_{Control\_HighImportance}$ = 2.98, SD = 1.12, $M_{Restricted\_LowImportance}$ = 2.18, SD = 1.10, $M_{Control\_LowImportance}$ = 2.46, SD = 1.33), the results are as follows. The analysis again indicates that there is no significant moderation effect of the level of importance on attitude toward the brand ($F(1, 396) = .08$, $p = .77$, $\eta^2 < .001$). However, the main effect of the post's status remains significant ($F(1, 396) = 7.19$, $p = .008$, $\eta^2 = .02$). In

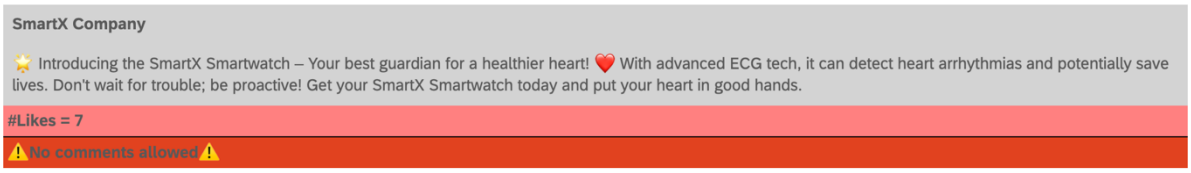
the high-importance conditions, restriction of the comments significantly influenced attitude toward the brand ($t = -2.13$, $p = .03$), while this effect became partially insignificant in the low importance conditions ($t = -1.72$, $p = .09$).

Discussion

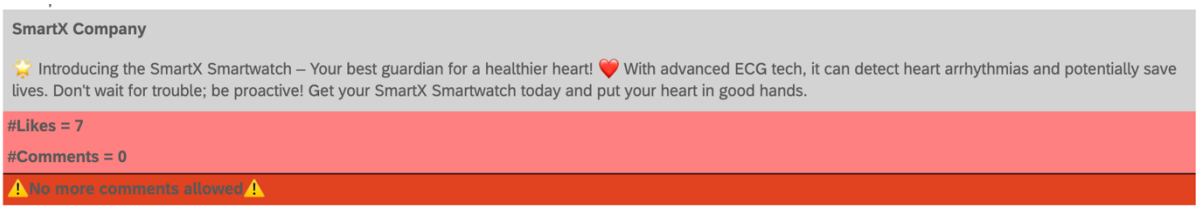
This study underscores the robustness of the effect between comment restriction and attitude toward the brand, regardless of the level of importance attributed to the post's topic. The findings from this study indicate that the drawbacks of limiting the comments are not confined solely to important and contentious topics.

Web Appendix 20. Stimuli Used in the Study 6

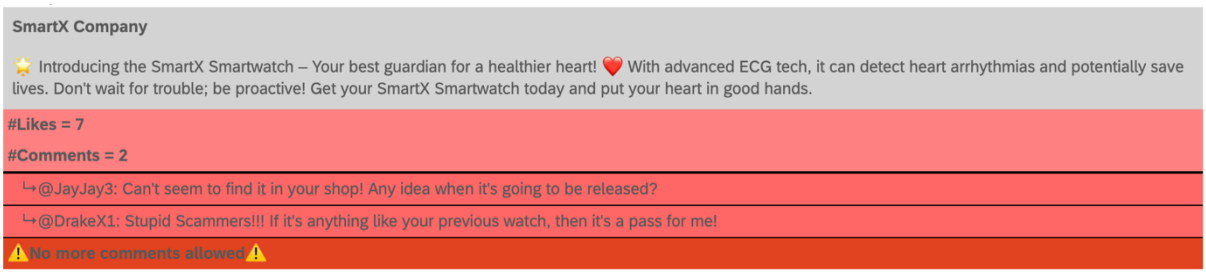
Restricted Before Posting



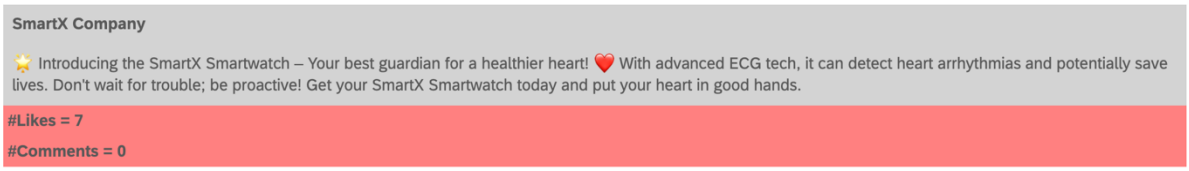
Restricted After Posting Without Any Comments



Restricted After Posting With Negative Comments



Control Without Any Comments



Control With Negative Comments