

How Does Social Media's Labeling Affect Users' Believability and Engagement? The Moderating Role of Regulatory Focus

Hui-Ying Han^a, Youngsok Bang^{b,*}

^a Master's Student, School of Business, Yonsei University, Korea

^b Associate Professor, School of Business, Yonsei University, Korea

ABSTRACT

In the wake of the COVID-19 pandemic, unsubstantiated information concerning vaccines and the coronavirus has proliferated on various social media platforms. Consequently, we have considered viable actions to mitigate the impact of such unverified content, enabling individuals to use social media platforms more effectively and minimize any ensuing confusion. Recent measures in this area have included YouTube's practice of labeling vaccine or corona videos as authoritative when emanating from reputable organizations and Twitter's practice of flagging vaccine-related content as potentially misleading or taken out of context. This study seeks to explore how such contrasting labeling practices influence users' believability and engagement differentially, while also examining the moderating impact of regulatory focus. The results indicate that authoritative labeling positively influenced users' believability and engagement, whereas misleading labeling adversely affected users' believability and engagement. Additionally, our findings revealed that authoritative labeling has a stronger impact on promotion-focused individuals, while misleading labeling has a more pronounced effect on prevention-focused individuals. Our findings offer insights into how social media platforms can design and present information to their users, taking into account their regulatory focus.

Keywords: Regulatory Focus, Labeling, Believability, Engagement, Fake News

1. Introduction

The rapid dissemination of misinformation through social media has precipitated adverse social consequences, particularly as pertains to the proliferation of inaccurate information regarding

COVID-19. The gravity of this issue has been widely recognized, as it has the potential to incite social unrest and hinder the efficient distribution of vaccines (e.g., Bin Naeem and Kamel Boulos, 2021; Enders et al., 2020; Pennycook et al., 2020a).

Various social media platforms have taken meas-

*Corresponding Author. E-mail: yb@yonsei.ac.kr

ures to counter the spread of COVID-19 misinformation. The White House has identified YouTube as a major contributor to the proliferation of COVID-19 vaccine misinformation, highlighting the platform's inadequacies in preventing such falsehoods from spreading. In a bid to improve its transparency, YouTube implemented a set of measures in July 2021, including displaying more trustworthy health information and attaching labels that viewers can click for more information. Likewise, Twitter has taken steps to address COVID-19 vaccine misinformation through two-pronged strategies: removing misleading information that poses a significant risk and labeling content that is either misleading or taken out of context. Misleading content that could be removed includes those that propagate false claims such as COVID-19 vaccines being part of a "deliberate conspiracy" or those that falsely contend that COVID-19 is a hoax and vaccines are unnecessary <Appendix A>.

Given the various recent labeling practices on platforms aimed at preventing the spread of fake information, researchers are now focusing their attention on the effectiveness of these efforts (Kim et al., 2019; Moravec et al., 2020, 2022). Notably, Mena (2020) demonstrates that employing flags for deceptive news has the potential to address the spread of misleading information on social media platforms like Facebook. Pennycook et al. (2020b) find spillover effects of content labeling where including notifications with some false news headlines increases the perceived accuracy of headlines that lack notifications. Nonetheless, these studies also indicate that a sole labeling approach cannot serve as a panacea for curbing the dissemination of misinformation (Morrow et al., 2022). Furthermore, no in-depth examination of the impact of YouTube or Twitter labeling measures has been undertaken.

Also, prior studies indicate that the effectiveness of content labeling can vary based on user heterogeneities. For instance, in reaction to warning messages on violent TV programs, it was observed that individuals under the age of 17 were more drawn to such programs compared to older age groups (Bushman, 2006). Alongside this demographic diversity, research findings also indicate that the effectiveness of labeling varies based on individuals' beliefs or preferences. An experiment shows that partisans evaluate news stories for their trustworthiness based on whether the news source (which may work as a content labeling) is perceived as supportive or antagonistic to their political affiliations (Knight Foundation and Gallup, 2018).

Drawing from the results indicating the influence of individual beliefs and preferences on the effectiveness of labeling, we can hypothesize that the alignment between a person's value system and external stimuli might play a fundamental role. Accordingly, this study posits that the impact of labeling as an external stimulus will vary depending on the regulatory fit with the user and delves into this effect more comprehensively.

Regulatory fit pertains to the perceived harmony between a person's orientation towards a goal and the means employed to attain it. This phenomenon does not necessarily influence the pleasurable aspect of an object or event but rather shapes a person's confidence in their reaction to it. The concept of regulatory fit posits that the alignment between goal orientation and the method employed to pursue it induces a state of regulatory fit, which amplifies the sense of appropriateness regarding goal pursuit and promotes task engagement (Higgins, 2005; Higgins et al., 2001). Regulatory fit can heighten various responses, such as the perceived value of a chosen object, persuasive communication, and job

satisfaction.

The primary objective of this study is to assess the efficacy of two novel labeling methods, namely authoritative and misleading labeling, on user credibility and engagement. The study focuses on users' regulatory focus and aims to examine the influence of this variable on the effectiveness of labeling practices. The results of the investigation indicate that authoritative labeling is more effective than the control condition in enhancing users' believability and engagement, with promotion-focused users responding more favorably than prevention-focused users. Conversely, misleading labeling diminishes users' believability and engagement in comparison to the control condition, with prevention-focused users exhibiting more adverse reactions than promotion-focused users. The study further demonstrates that users' regulatory focus can be rapidly manipulated through message framing.

II. Literature Review

2.1. Fake News, Misinformation, Unproven Information

The term "fake news" denotes media content that is either deliberately falsified or manipulated, encompassing both satirical websites and established outlets such as BBC News and Sky News. While the term has gained widespread attention in recent times, it was first coined in 2003 by American journalist David Sirota, who described it as an intentional attempt to deceive. The Oxford English Dictionary defines fake news as "fabricated or distorted information presented in the form of news reporting," and it is not limited to conventional journalism but may also appear on various platforms, including social

media.

The issue of media fake news is not novel and can be traced back to the early stages of writing systems. However, its prevalence has surged in the era of social media. As of December 2016, Facebook, the most popular social media platform, reported over 1.23 billion daily active users (Swant, 2017), signifying an expanding potential audience for fake news. The original intent behind social media was not to disseminate false information. Nonetheless, research conducted by Nelson and Taneja (2018) suggests that a small and unrepresentative group of frequent internet users constitute the fake news audience. Moreover, social networking sites play an outsized role in driving traffic to fake news sources. A study conducted by Grinberg et al. (2019) analyzed the exposure of registered voters to fake news on Twitter and revealed that engagement with fake news sources was markedly concentrated. Only 1% of individuals were responsible for 80% of exposures to false news sources, while 0.1% were accountable for over 80% of shared fake news sources.

The reasons behind the sharing of misinformation on social media by users are multifaceted. Taiwar et al. (2019) found that sharing of false information was positively associated with online trust, self-disclosure, fear of missing out (FoMO), and fatigue, while online social comparison was negatively correlated with it. Additionally, an increasing number of adults are utilizing social media as their primary source of information (Gottfried and Shearer, 2016). Previous research has indicated that users often encounter challenges in distinguishing between authentic and fake information (Silverman, 2016), with confirmation bias being a significant factor (Nickerson, 1998).

Fake news often employs multimedia content to mislead readers and can significantly impact public

events. The detection of fake news related to recent events poses a challenging task in identifying fake news on social media. Current techniques encounter difficulties in addressing this issue since they rely on event-specific properties that do not apply to future events. The automatic detection of fake news has garnered significant interest from both scholars and the general public, given the widespread influence of social media. Existing detection methods employ machine learning algorithms that leverage diverse news properties. Nevertheless, the information necessary for the early detection of false news is often unavailable or insufficient, posing a significant limitation of such systems. In response, Liu and Wu (2018) proposed a novel model for the early detection of fake news on social media that categorizes news dissemination channels. The proposed model achieved an accuracy of 85% and 92% on Twitter and Sina Weibo, respectively, in detecting fake news within five minutes of its initiation, surpassing state-of-the-art baselines on three real-world datasets. Similarly, Wang et al. (2018) proposed a distinct model named Event Adversarial Neural Network (EANN) that enhances the detection of fake news related to recently received events. EANN is comprised of three primary components: a multi-modal feature extractor, a fake news detector, and an event discriminator. The experimental results demonstrated that the proposed EANN model can learn transferable feature representations and outperform state-of-the-art approaches.

In response to the proliferation of fake news, numerous fact-checking organizations have emerged, which evaluate the accuracy of each news report (Graves, 2016; Lowrey, 2017; Wintersieck, 2017). One method for supplementing or complementing the fact-checking of articles or authors/sites is to rate the credibility of the original sources. However, re-

search on authoritative labeling and deceptive labeling based on fact-checking is scarce, leading to the development of two labeling measures that can provide more context and guidance for fact-checking projects. Previous research has demonstrated that individual fact-checking initiatives impact their perceived credibility (Wintersieck, 2017).

2.2. Social Media Believability and Engagement

The Internet serves various purposes for individuals, including goal-oriented activities and hedonic pursuits (Zhou et al., 2011). Although social media platforms are employed by many for practical purposes such as information retrieval or professional communication (Johnson and Kaye, 2015), they are primarily utilized for hedonic aims such as entertainment and social interaction (Harsanyi, 1977). Consequently, individuals with hedonic mindsets, who associate social media use with pleasure and enjoyment, may exhibit less inclination to scrutinize information critically relative to those with utilitarian mindsets (Hirschman and Holbrook, 1982).

The concept of believability has been the subject of inquiry since ancient times, with philosophers such as Plato and Aristotle (circa 4th century BC) exploring its implications. In modern times, the domains of communication, psychology, and social science have examined believability, which is defined as trustworthiness, credibility, perceived reliability, accuracy, and knowledge, and is associated with numerous other connotations contingent on the context. In the study of credibility and information credibility assessment, mass and online communication have rapidly supplanted conventional communication (Salwen et al., 2004). While interpersonal and persuasive communication serves as the cornerstone

of traditional “offline” environments, “online” environments are distinguished by the absence of direct familiarity between individuals and traditional intermediaries in disseminating authenticated information.

Believability consists of multiple interrelated elements, such as the information source, the message’s content and structure, and the medium employed to convey the information (Metzger et al., 2003). The perception of the information source and message content may be susceptible to influence, particularly by the medium employed to disseminate the information (Metzger et al., 2003; Salwen et al., 2008). As a result, current research is focused on how new digital media affect the evaluation of information believability, taking into consideration the distinct attributes of online environments (Metzger et al., 2003).

Evaluating information believability on social media platforms entails scrutinizing user attributes, user-generated content, and social connections linking users and other entities. These features can be classified into multiple types: (i) linguistic features on the textual content of social media posts; (ii) meta-data features, concerning user behavior on social media, including reviews or tweets; (iii) behavioral features; (iv) product-based features, applicable in scenarios where goods or services are evaluated; and (v) social features, which are especially pertinent as they employ the social network structure that interconnects users and entities on social media platforms (Heydari et al., 2015). If available, user profiles can also be employed to extract user-related features.

Consumer engagement can be classified into three distinct types: cognitive, affective, and behavioral (Dessart, 2017; Dessart et al., 2015; Muntinga et al., 2011). According to Dessart (2017), cognitive engagement is analogous to the cognitive processes involved in focusing attention and assimilating information.

Affective engagement pertains to emotional responses such as enthusiasm and enjoyment. The term “behavioral engagement” refers to actions that demonstrate the engagement notion, such as sharing, learning, and affirming behaviors. This research aims to expand and advance the behavioral perspective, which is more useful for the analytic measures employed to evaluate social media engagement success (Pentina et al., 2018).

Assessing customer responses to posted content is a crucial responsibility of destination marketing since active social media interaction expands the brand reach and subsequently impacts economic success (Kim and Kim, 2020; Nezakati et al., 2015; Sparks et al., 2013). Social media engagement is a type of behavior driven by emotional responses to a post (Zaidi et al., 2020). The number of likes, comments, and shares, as well as engagement rates, can serve as indicators of behavioral engagement (Demmers et al., 2020; Lalicic et al., 2020; Leung, 2019; Solem and Pedersen, 2016). While these measures may translate slowly into improved sales, booking rates, or expenditure, they reveal consumers’ willingness to engage with a brand and their attitudes toward the destination content (Demmers et al., 2020). All three response behaviors - likes, comments, and shares - reflect varying degrees of psychological and behavioral involvement. Clicking “like” is the least time-consuming response, while commenting and sharing a social media post demonstrate higher involvement (Solvoll and Larsson, 2020).

2.3. Regulatory Focus Theory

Regulatory focus theory, also known as regulatory orientation theory, explains how individuals pursue their goals while adhering to their values and beliefs (Avnet and Higgins, 2003). This theory is based on

the core principle that people seek pleasure while avoiding pain and aims to maintain an individual's regulatory fit. Regulatory focus refers to how individuals approach pleasure and avoid pain, primarily by pursuing desired end-states and adopting motivational strategies to transition from the present state to the desired end-state. This concept distinguishes between two types of promotion: hope-focused and accomplishment-focused, also referred to as gains. In this view, higher-order gains such as advancement and achievement hold greater importance (Higgins et al., 2001). Another focus area is prevention focus, also known as non-loss orientation, which emphasizes safety and responsibility. This focus prioritizes security and safety by complying with rules and regulations (Higgins et al., 2001).

Following regulatory focus theory, a person's orientation towards regulation affects the outcomes they experience while making decisions and the methods they use to reach their goals. Nevertheless, regulatory orientation is not fixed, and while some people may naturally prefer promotion or prevention, this preference may not always be appropriate. Furthermore, a specific regulatory focus can be induced, and the benefits of engagement and goal attainment can be positive or negative, depending on whether they align with the individual's regulatory orientation (Avnet and Higgins, 2003). Regulatory fit can enhance value by aligning commitment with one of the regulatory orientations. Decision-making and goal attainment are seen as activities, and high engagement can intensify emotions and values related to the activity. The process used to arrive at the result affects the level of satisfaction, which has significant implications for increasing the value of life. For example, when all parties involved in a conflict experience a "fit," they are more likely to be satisfied with and committed to the resolution. Individuals must feel

satisfied with their actions and "feel right" about them to value their own lives (Higgins, 2005). If their goal is "non-fit" or unsatisfactory, they are unlikely to succeed in achieving it.

In various fields, regulatory focus theory has been widely used to study human behavior. For instance, Petrou et al. (2018) found that effective change communication is associated with an increase in job-crafting behaviors among promotion-focused employees, whereas ineffective change communication is associated with an increase in job-crafting behaviors among prevention-focused employees. Similarly, Lin et al. (2018) found that customer happiness has a greater impact on attitudinal and behavioral loyalty than promotion focus, and promotion focus and cognitive dissonance have significant effects on consumer satisfaction. In the context of employee creativity, Geng et al. (2018) found that promotion focus is positively related to frontline employee creativity, while prevention focus is negatively related to it. Moreover, the relationship between prevention focus and entrepreneurial improvisation is significantly moderated by environmental volatility.

Regarding consumer regulatory focus, Song and Qu (2019) found that individuals more interested in promotion are more likely to perceive both hedonistic and utilitarian values and feel good while eating at Asian restaurants, whereas prevention focus had a significant impact on unpleasant emotions but was not significantly related to either hedonic or utilitarian values in itself. Studies on message framing have found that promotional messages work better in harsh environments when they align with a "how" construal perspective (Cai and Leung, 2020). In places with moderate pandemics, promotion-framed communications are more persuasive in speculation with a "why" construal attitude. Tretter and Diefenbach (2021) suggested that the introduction of a dominant

prevention focus (vs. promotion focus) leads to a shift in choice towards leaner communication devices and channels that create a larger subjective buffer between sender and receiver, such as text messaging over calling. In addition, Mühlberger et al. (2022) found that promotion focus was associated with fluency, and prevention focus with elaborated ideas. This effect was more pronounced when prevention also scored highly on avoidance motivation. Furthermore, prevention-focused individuals who performed well on avoidance reported feeling happier and receiving more autonomy support.

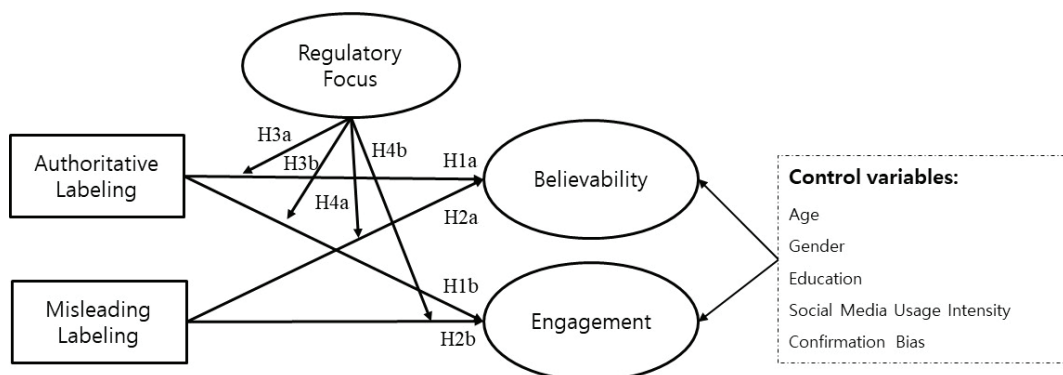
In various studies, regulatory focus has been analyzed as a moderator of different relationships. In one study, the relationship between career plateau and job burnout was examined, and the moderating effect of regulatory focus was investigated. The study found that career plateau had a significant direct impact on job burnout, and this relationship persisted even after accounting for the interaction with regulatory focus. Additionally, the results showed that promotion focus had a harmful moderating effect, whereas prevention focus did not affect the relationship between career plateau and job burnout (Kwon, 2022). Another study investigated the moderating role of regulatory focus in the relationship between

cause-related marketing and consumers' extra-role behavior. The findings indicated that customer promotion focus significantly moderated the beneficial impact of altruistic cause-related marketing on customer extra-role behavior, as well as the negative impact of egoistic cause-related marketing on customer extra-role behavior (Hui et al., 2022).

III. Research Model

Our main research model is shown below in <Figure 1>.

Authoritative labeling provides information about the source of the information, which can help users assess its credibility. When a piece of information is labeled as coming from an authoritative source, it signals to users that the information has been verified by a trustworthy and knowledgeable entity. This, in turn, can increase users' confidence in the information and its credibility. Additionally, authoritative labeling may also serve as a cue to users that the information is more important or relevant, which can further enhance its believability. Overall, authoritative labeling can positively affect believability by providing users with valuable information about the



<Figure 1> Research Model

source of the information that can help them make more informed judgments about its credibility.

Authoritative labeling can positively affect engagement because it can enhance the perceived credibility of the information source, message content, and medium, which may increase users' motivation to engage with the information. When users perceive a source as authoritative, they may be more likely to engage with the content, such as by sharing or commenting on the post, because they trust the information and believe it is worth sharing with others. Additionally, authoritative labeling can increase the salience of the message, making it more visible and thus more likely to attract user attention and engagement. Finally, authoritative labeling can serve as a signal of quality and value, which can increase user interest in the content and lead to greater engagement. Overall, authoritative labeling can enhance the believability and value of the message, increasing engagement in the process.

Prior studies also provide empirical evidence that authoritative labeling may enhance both believability and engagement. Arnold et al. (2021) show that putting authoritative source labels (like BBC) has the potential to reduce belief in foreign misinformation and temper the propensity to share. In a similar vein, despite awareness of the potential content manipulation, government-controlled channels like Russia Today can still wield influence over the opinions and perceptions of viewers (Nassetta and Gross, 2020).

Thereby, we propose that:

H1a: Authoritative labeling positively affects believability.

H1b: Authoritative labeling positively affects engagement.

In contrast, misleading labeling involves providing false or misleading information that deceives users. This can lead to a reduction in the trustworthiness, reliability, and accuracy of the information source. When users encounter misleading labeling, they may question the credibility of the information source and the information provided, leading to a decrease in their level of trust and engagement with the source. Moreover, users may feel misled or deceived, leading to negative emotions and a reluctance to engage with the source in the future. In this way, misleading labeling can harm the believability of the information source and reduce user engagement.

Regarding this, a recent experiment illustrates that the utilization of flags to mark deceptive news diminishes its credibility and reduces intentions to share such content (Mena, 2020). Accordingly, we propose that:

H2a: Misleading labeling negatively affects believability.

H2b: Misleading labeling negatively affects engagement.

Promotion-focused individuals have a regulatory focus on achieving gains, advancement, and accomplishment, which may result in a greater willingness to trust and accept authoritative labels. This is because promotion-focused individuals are more inclined to seek out positive outcomes and are more responsive to opportunities to achieve them. In contrast, prevention-focused individuals tend to focus on safety and responsibility, which may result in a greater skepticism toward information that appears to be overly certain or authoritative. This could explain why they are less likely to accept authoritative labeling.

Promotion-focused individuals may naturally in-

cline towards authoritative labels that assure accurate information, as they perceive these labels as opportunities for gains. Driven by the desire for rewards and positive outcomes, they are more likely to affiliate themselves with sources that promise reliable information. In the context of our experimental focus on COVID-19 misinformation, where accurate knowledge directly correlates with improved health outcomes, promotion-focused individuals may exhibit heightened motivation to align with sources that offer a higher likelihood of providing correct information.

Thus, we propose that:

H3a: Promotion-focused individuals are more susceptible to the influence of authoritative labeling compared to prevention-focused individuals concerning their believability.

H3b: Promotion-focused individuals are more susceptible to the influence of authoritative labeling compared to prevention-focused individuals concerning their engagement.

Individuals with a prevention focus are more cautious and risk-averse, leading them to be more likely to scrutinize information for potential threats or negative outcomes. This may result in prevention-focused individuals being more sensitive to misleading information and less likely to believe it, as it conflicts with their natural inclination to avoid negative consequences.

By way of example within the realm of COVID-19 misinformation, instances of misleading labeling can be elucidated. In these cases, the labeling does not explicitly ascertain the veracity of the information provided, concurrently acknowledging the plausible inaccuracy inherent to the information. Such instances of labeling predominantly function to attenu-

ate potential negative consequences, a convergence that resonates more strongly with individuals inclined toward prevention-focused orientations, wherein the emphasis lies on averting losses rather than pursuing gains. Thus, we propose that:

H4a: Prevention-focused individuals are more susceptible to the influence of misleading labeling compared to promotion-focused individuals concerning their believability.

H4b: Prevention-focused individuals are more susceptible to the influence of misleading labeling compared to promotion-focused individuals concerning their engagement.

Nonetheless, it is important to consider that the relationship between labelings, regulatory focus, and believability (engagement) may be impacted by various individual factors, including individual heterogeneities and biases. To address this issue, we take into account several variables, including age, gender, education level, social media usage intensity, and confirmation biases, when estimating the relationships between these constructs.

IV. Empirical Approach

4.1. Experiment

We conducted a vignette experiment to test our hypotheses, gathering data from 385 Xiaohongshu users through the Tencent Questionnaire platform. Following the removal of participants who did not notice the treatment (manipulation check), our final sample size was 325 participants, with an average age of 24 and 70% of participants identifying as female. We also assessed their frequency of

Xiaohongshu platform use, finding that 29% (95 participants) used the platform once a week or less, 26% (83 participants) used it more than once a week, and 45% (147 participants) used it more than once a day. <Table 1> shows participant demographics in our analysis.

We created a mock post titled “Be cautious about getting the COVID-19 vaccine booster shot!” (<Figure 2>). Participants were randomly assigned to Treatment 1 and 2 groups. Treatment 1 involved posts with authoritative labeling (<Figure 2B>), which included a tag linking to reliable sources recommended by the National Academy of Medical Science, while Treatment 2 had misleading labeling (<Figure 2C>), with the tag “Get the facts about the information.” After the treatment, participants provided their believability and engagement intentions again. The experiment was designed using a within-subject and between-subject design adapted from Kim and Dennis (2019), controlling for potential confounding factors such as age, gender, education, social media usage intensity, and confirmation biases.

<Table 1> Participant demographics (n=325)

	Mean	Standard Deviation
Age	23.5	5.5
	Frequency	Percentage
Gender		
- Female	226	69.5%
- Male	99	30.5%
Education		
- Less than a high school degree	14	4.3%
- High school degree or equivalent	67	20.6%
- Bachelor's (or associate) degree or equivalent	233	71.7%
- Graduate degree	11	3.4%
SM Usage Intensity		
- Once a week or less	95	29.2%
- More than once a week, but less than once a day	83	25.5%
- More than once a day	147	45.2%



<Figure 2A> Messages without Labeling (Baseline)



<Figure 2B> Messages with Authoritative Labeling (Treatment 1)



<Figure 2C> Messages with Misleading Labeling (Treatment 2)

4.2. Measures and Variables

We utilized three 7-point items to measure our focal variables, *Believability* and *Engagement* (Kim and Denis, 2019). *Believability* was assessed through items such as “How believable do you find this article?”, “How truthful do you find this article?”, and

“How credible do you find this article?,” while *Engagement* was measured through items such as “How likely are you to ‘like’ this article?”, “How likely are you to leave a comment on this article?”, and “How likely are you to share this article?”. To assess regulatory focus, we used four items based on a 5-point scale, adapted from Higgins et al. (2001).

<Table 2> Descriptive Statistics of the Variables

Variables	Mean	SD	Minimum	Maximum
BLV _{before}	3.61	1.53	1.00	7.00
BLV _{after}	3.52	1.71	1.00	7.00
ENG _{before}	2.73	1.44	1.00	7.00
ENG _{after}	3.04	1.50	1.00	7.00
RF	2.83	0.88	1.00	5.00
CB	3.30	7.53	-21.00	21.00

Note: BLV_{before, after} (*Believability*) and ENG_{before, after} (*Engagement*) are the average values of items corresponding to the construct before and after the treatment. RF (*Regulatory Focus*) is the average value of four items measuring *Regulatory Focus*. CB (*Confirmation Bias*) is the multiplication of the post’s importance to the participant and the participant’s stance.

<Table 3> Composite Reliability (CR), Cronbach’s α , Correlation, and AVE

Variables	CR	Cronbach’s α	BLV _{before}	BLV _{after}	ENG _{before}	ENG _{after}	RF
BLV _{before}	0.969	0.968	0.955				
BLV _{after}	0.964	0.964	0.546	0.948			
ENG _{before}	0.923	0.922	0.787	0.465	0.895		
ENG _{after}	0.943	0.942	0.647	0.639	0.872	0.920	
RF	0.803	0.796	0.364	0.201	0.323	0.270	0.728

Note: The values on the diagonal of the correlation matrix are the square roots of the Average Variance Extracted (AVE).

The full list of items can be found in <Appendix B>.

To control for potential confounding factors, we incorporated several control variables into the study, such as age, gender, education, social media usage intensity, and confirmation bias. Prior research suggests that individuals are more inclined to believe information that aligns with their existing beliefs, with confirmation bias being a significant contributing factor (Devine et al., 1990; Koriat et al., 1980; Nickerson, 1998). Following Kim and Denis (2019), we controlled for the confirmation bias, which was measured by multiplying the posting’s importance to the participant (Do you find the issue described in the posting important? 1 = not at all, 7 = extremely) by the participant’s stance (−3 = extremely negative to +3 = extremely positive), rang-

ing from −21 to +21. <Table 2> shows descriptive statistics of the variables used in our analysis.

Convergent validity is generally assessed based on Composite Reliability (CR) and Average Variance Extracted (AVE) (Xu et al., 2010). The Composite Reliability surpasses the criterion of 0.7 as suggested by Nunnally (1978) (<Table 3>). The AVE for each construct also exceeds the general threshold of 0.5 (Chin, 1998), indicating overall convergent validity. Cronbach’s alpha values also appeared to be above 0.7. For the assessment of discriminant validity, a comparison of the square root of the AVE for each construct with the correlation coefficients with other constructs showed that the variance shared by each construct with its related measurement items was greater than the variance shared with other constructs (<Table 3>).

V. Results

We evaluated the participants' *believability* and *engagement* both prior to and following the treatment to observe any modifications in their responses. To analyze these changes, we utilized the repeated measures analysis of covariance (ANCOVA). Repeated measures regression such as ANCOVA is a statistical technique used to analyze repeated measures data, where multiple measurements are taken on the same subjects over time or under different conditions. In our case, the dependent variables (*Believability* and *Engagement*) are measured before the experiment and again after the experiment. This methodology is particularly suitable for capturing such within-subject changes while controlling for individual differences (Lorch and Myers, 1990). It has already been used in multiple studies with similar experimental settings (e.g., Cazier et al., 2017; Chow and Luk, 2006; Kleber et al., 2016; Toreini et al., 2022).

Our specification is:

$$Y_{ijt} = \beta_0 + \beta_1 T_{ij} + \beta_2 Y_{ijt} + \beta_3 RF_i^* T_{ij} + Control_i + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2) \quad (1)$$

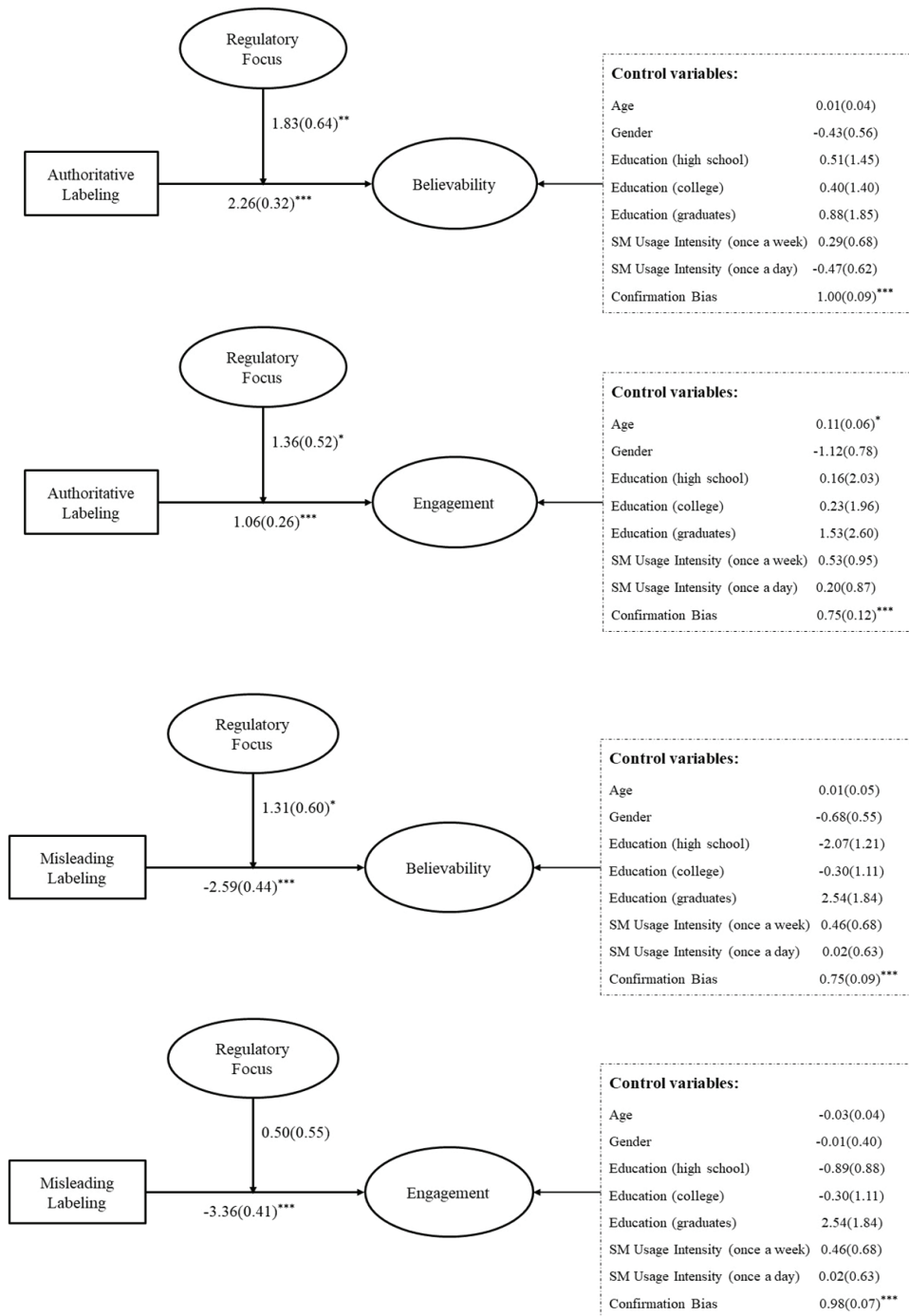
where $Y_{ij(\alpha)}$ is the *Believability* or *Engagement* of individual i before (after) treatment j , T_{ij} is the treatment j of individual i (authoritative labeling vs. misleading labeling), RF_i is the regulatory focus score of individual i , $Control_i$ is a set of control variables, and ε_{ij} is the error term.

<Table 4> and <Figure 3> present the results of our analysis. The coefficient for *authoritative labeling* is positive and statistically significant for both *believability* and *engagement*, thus providing support for H1a and H1b. The interaction term between *authoritative labeling* and *regulatory focus* also has a positive and significant coefficient for both *believability* and *engagement*, which supports H3a and H3b. Our control variables show that confirmation biases have a significant effect on both *believability* and *engagement*, and *age* is positively associated with *engagement*.

In contrast, the coefficient for *misleading labeling* is negative and statistically significant for both *believability* and *engagement*, providing evidence for H2a and H2b. The interaction term between *misleading labeling* and *regulatory focus* has a negative and significant coefficient for *believability*, but it is not significant for *engagement*, thus providing support only

<Table 4> Repeated Measures Regression Result

Authoritative labeling (Treatment 1)				
Path	Path coefficient	Standardized Error	p-value	Hypothesis
H1a: <i>Labeling</i> → <i>Believability</i>	2.26	0.32	< 0.001	Supported
H1b: <i>Labeling</i> → <i>Engagement</i>	1.06	0.26	< 0.001	Supported
H3a: <i>Labeling*</i> <i>Regulatory focus</i>	1.83	0.64	< 0.01	Supported
H3b: <i>Labeling*</i> <i>Regulatory focus</i>	1.36	0.52	< 0.05	Supported
Misleading labeling (Treatment 2)				
H2a: <i>Labeling</i> → <i>Believability</i>	-2.59	0.44	< 0.001	Supported
H2b: <i>Labeling</i> → <i>Engagement</i>	-3.36	0.41	< 0.001	Supported
H4a: <i>Labeling*</i> <i>Regulatory focus</i>	1.31	0.60	< 0.05	Supported
H4b: <i>Labeling*</i> <i>Regulatory focus</i>	0.50	0.55	> 0.05	Rejected



<Figure 3> Repeated Measures Regression Result

Note: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors are in parentheses.

for H4a.

In our set of hypotheses, Hypothesis 4b did not receive empirical support. The construct of Engagement is intricately linked to tangible behavioral modifications, necessitating a greater magnitude of impact to induce change compared to the construct of Believability. As evidenced in <Table 4>, the baseline intention for engaging in actions such as ‘liking,’ ‘sharing,’ or ‘commenting’ was inherently low; hence, any incremental shifts in intentionality were not statistically significant. Particularly, authoritative labeling, which serves to validate the content’s veracity and thereby its value, had a more pronounced effect compared to misleading labeling. The latter assumes a preventive function by signaling the potential inaccuracy of the content, thereby mitigating prospective losses, which could explain its diminished impact.

Our control variables also show that confirmation biases have a significant impact on both *believability* and *engagement*. Other individual characteristics were not significant. Methodologically, this could be because Y_{j0} in Equation 1 controls for the base level. Since the repeated measures ANCOVA examines the differences before and after the experiment, individual characteristics related to levels rather than differences may become statistically insignificant.

VI. Discussion and Conclusion

Amid the COVID-19 pandemic, social media has seen an influx of unverified information regarding vaccines and the virus. To address this issue and promote effective use of social media platforms, measures have been implemented to combat the spread of misinformation. One such measure involves labeling vaccine and coronavirus videos on YouTube as

authoritative when coming from reputable sources, while Twitter flags vaccine-related content as potentially misleading or taken out of context. This study aims to investigate how these contrasting labeling practices differentially influence users’ credibility assessments and behaviors and the moderating impact of regulatory focus.

The results suggest that both authoritative and misleading labeling measures significantly impact users’ believability and behavior. Authoritative (misleading) labeling enhances (reduces) users’ believability and engagement, with a stronger effect on individuals with a promotion-focused (prevention-focused) orientation. This finding implies that authoritative (misleading) labeling is especially effective for users who are more likely to respond to positive (negative) information.

From a theoretical perspective, our research sheds light on the interplay between individuals’ regulatory focus orientations and their responses to labeled information. This enriches the existing literature by revealing the nuanced ways in which psychological factors influence the processing of labeled content. By identifying the distinct pathways through which promotion-focused and prevention-focused individuals interpret and engage with labeled information, we contribute to a deeper comprehension of cognitive processes and decision-making mechanisms.

Existing research on the labeling effect to prevent the spread of fake news focuses on the average treatment effect of labeling, without considering individual heterogeneity (Morrow et al., 2022). As media content and communication are becoming increasingly hyper-personalized, it is technically feasible to provide such labeling in a customized manner for each individual. In this context, this study is the first to show that the effect of labeling varies

depending on an individual's regulatory focus, providing important implications in the setting of personalized media and message delivery. Specifically, the study's results highlight the importance of an individual's regulatory focus in their response to different labeling measures. Authoritative labeling is more effective for promotion-focused individuals, while misleading labeling has a more detrimental impact on prevention-focused individuals. Therefore, regulatory focus should be taken into account when designing and implementing labeling regulations on social media platforms to ensure their effectiveness for all users.

What should be done if it is difficult to determine the user's regulatory focus? Prior studies argue that the use of situational factors such as message framing can trigger a temporary shift in an individual's regulatory focus orientation (Florack et al., 2013). In other words, an individual's regulatory focus in a given situation can change temporarily based on the situational factors present. The findings of this study, which show that the same message can have different effects on believability and engagement depending on the user's regulatory focus, provide important practical implications related to how social media platforms can design and present information to their users. In light of the study's results, it is crucial to acknowledge an individual's regulatory focus as a situational feature and the potential effectiveness of situational framing in temporarily activating regu-

latory focus. These findings have practical implications for social media platforms in designing effective labeling regulations that can assist users in identifying trustworthy information.

In addition, the study recognizes the necessity for further research on topics beyond medical information to enhance our understanding of the effectiveness of labeling approaches. Our experiment is limited to a single hypothetical social media context related to COVID-19. Future studies may employ diverse content tasks to improve the generalizability of the results and provide more extensive suggestions for research on fact-checking policies. Investigating the effects of labeling on other variables, such as actual behavior and decision-making, is also recommended. Nonetheless, the study is valuable in contributing to the field of social media and misinformation, as it sheds light on the effectiveness of labeling measures and the impact of regulatory focus.

Acknowledgments

This research was financially supported by the Yonsei Business Research Institute, the Yonsei Signature Research Cluster Program of 2023-22-0014 and the BK21 FOUR (Fostering Outstanding Universities for Research) in 2022.

<References>

- [1] Arnold, J. R., Reckendorf, A., and Wintersieck, A. L. (2021). Source alerts can reduce the harms of foreign disinformation. *Harvard Kennedy School Misinformation Review*, 1, 1-33. <https://doi.org/10.37016/mr-2020-68>
- [2] Avnet, T., and Higgins, E. T. (2003). Locomotion, assessment, and regulatory fit: Value transfer from "how" to "what". *Journal of Experimental Social Psychology*, 39(5), 525-530.
- [3] Bollen, K. A., and Pearl, J. (2013). Eight myths about

- causality and structural equation models. In *Handbook of Causal Analysis for Social Research* (pp. 301-328). Dordrecht: Springer Netherlands.
- [4] Bin Naeem, S., and Kamel Boulos, M. N. (2021). COVID-19 misinformation online and health literacy: A brief overview. *International Journal of Environmental Research and Public Health*, *18*(15), 8091.
- [5] Bushman, B. J. (2006). Effects of warning and information labels on attraction to television violence in viewers of different ages. *Journal of Applied Social Psychology*, *36*(9), 2073-2078. <https://doi.org/10.1111/j.0021-9029.2006.00094.x>
- [6] Cai, R., and Leung, X. Y. (2020). Mindset matters in purchasing online food deliveries during the pandemic: The application of construal level and regulatory focus theories. *International Journal of Hospitality Management*, *91*, 102677. <https://doi.org/10.1016/j.ijhm.2020.102677>
- [7] Cazier, J., Shao, B., and Louis, R. S. (2017). Value congruence, trust, and their effects on purchase intention and reservation price. *ACM Transactions on Management Information Systems*, *8*(4), 1-28. <https://doi.org/10.1145/3110939>
- [8] Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295-336). Mahwah, NJ: Lawrence Erlbaum Associates.
- [9] Chow, C. W., and Luk, C. L. (2006). Effects of comparative advertising in high-and low-cognitive elaboration conditions. *Journal of Advertising*, *35*(2), 55-67. <https://doi.org/10.1080/00913367.2006.10639229>
- [10] Demmers, J., Weltevreden, J. W., and van Dolen, W. M. (2020). Consumer engagement with brand posts on social media in consecutive stages of the customer journey. *International Journal of Electronic Commerce*, *24*(1), 53-77. <https://doi.org/10.1080/10864415.2019.1683701>
- [11] Dessart, L. (2017). Social media engagement: A model of antecedents and relational outcomes. *Journal of Marketing Management*, *33*(5-6), 375-399. <https://doi.org/10.1080/0267257X.2017.1302975>
- [12] Dessart, L., Veloutsou, C., and Morgan-Thomas, A. (2015). Consumer engagement in online brand communities: A social media perspective. *Journal of Product & Brand Management*, *24*(1), 28-42. <https://doi.org/10.1108/JPB-06-2014-0635>
- [13] Devine, P. G., Hirt, E. R., and Gehrke, E. M. (1990). Diagnostic and confirmation strategies in trait hypothesis testing. *Journal of Personality and Social Psychology*, *58*(6), 952-963. <https://doi.org/10.1037/0022-3514.58.6.952>
- [14] Enders, A. M., Uscinski, J. E., Klofstad, C., and Stoler, J. (2020). The different forms of COVID-19 misinformation and their consequences. *The Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-48>
- [15] Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. (2012). *Applied Longitudinal Analysis*. John Wiley & Sons.
- [16] Florack, A., Keller, J., and Palcu, J. (2013). Regulatory focus in economic contexts. *Journal of Economic Psychology*, *38*, 127-137. <https://doi.org/10.1016/j.joep.2013.06.001>
- [17] Geng, Z., Li, C., Bi, K., Zheng, H., and Yang, X. (2018). Motivating service employee creativity: Regulatory focus and emotional labour. *Journal of Service Theory and Practice*, *28*(2), 228-249. <https://doi.org/10.1108/JSTP-11-2016-0214>
- [18] Gottfried, J., and Shearer, E. (2016). *News Use Across Social Media Platforms*. Pew Research Center.
- [19] Graves, L. (2016). *Deciding what's true: The rise of political fact-checking in American journalism*. Columbia University Press.
- [20] Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., and Lazer, D. (2019). Fake news on Twitter during the 2016 US presidential election. *Science*, *363*(6425), 374-378.
- [21] Harsanyi, J. C. (1977). Morality and the theory of rational behavior. *Social Research*, 623-656. <https://doi.org/10.1093/acprof:oso/9780199253609.003.0013>
- [22] Heydari, A., Tavakoli, M. A., Salim, N., and Heydari,

- Z. (2015). Detection of review spam: A survey. *Expert Systems with Applications*, 42(7), 3634-3642.
- [23] Higgins, E. T. (2005). Value from regulatory fit. *Current Directions in Psychological Science*, 14(4), 209-213. <https://doi.org/10.1111/j.0963-7214.2005.00366.x>
- [24] Higgins, E. T., Friedman, R. S., Harlow, R. E., Idson, L. C., Ayduk, O. N., and Taylor, A. (2001). Achievement orientations from subjective histories of success: Promotion pride versus prevention pride. *European Journal of Social Psychology*, 31(1), 3-23. <https://doi.org/10.1002/ejsp.27>
- [25] Hirschman, E. C., and Holbrook, M. B. (1982). Hedonic consumption: Emerging concepts, methods and propositions. *Journal of Marketing*, 46(3), 92-101.
- [26] Johnson, T. J., and Kaye, B. K. (2015). Reasons to believe: Influence of credibility on motivations for using social networks. *Computers in Human Behavior*, 50, 544-555. <https://doi.org/10.1016/j.chb.2015.04.002>
- [27] Kim, A., and Dennis, A. R. (2019). Says who? The effects of presentation format and source rating on fake news in social media. *MIS Quarterly*, 43(3), 1025-1039.
- [28] Kim, A., Moravec, P. L., and Dennis, A. R. (2019). Combating fake news on social media with source ratings: The effects of user and expert reputation ratings. *Journal of Management Information Systems*, 36(3), 931-968. <https://doi.org/10.1080/07421222.2019.1628921>
- [29] Kim, A., Moravec, P. L., and Dennis, A. R. (2019). Combating fake news on social media with source ratings: The effects of user and expert reputation ratings. *Journal of Management Information Systems*, 36(3), 931-968. <https://doi.org/10.1080/07421222.2019.1628921>
- [30] Kim, M., and Kim, J. (2020). Destination authenticity as a trigger of tourists' online engagement on social media. *Journal of Travel Research*, 59(7), 1238-1252. <https://doi.org/10.1177/0047287519878510>
- [31] Kleber, J., Florack, A., and Chladek, A. (2016). How to present donations: The moderating role of numeracy in cause-related marketing. *Journal of Consumer Marketing*, 33(3), 153-161.
- [32] Knight Foundation and Gallup. (2018). An online experimental platform to assess Trust in the Media Knight Foundation. Retrieved from https://knightfoundation.org/wp-content/uploads/2020/02/KnightFoundation_NewsLens1_Client_Report_070918_ab.pdf
- [33] Koriat, A., Lichtenstein, S., and Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 107-118. <https://doi.org/10.1037/0278-7393.6.2.107>
- [34] Kwon, J. E. (2022). The Impact of career plateau on job burnout in the COVID-19 pandemic: A moderating role of regulatory focus. *International Journal of Environmental Research and Public Health*, 19(3), 1087. <https://doi.org/10.3390/ijerph19031087>
- [35] Lalicic, L., Huertas, A., Moreno, A., and Jabreel, M. (2020). Emotional brand communication on Facebook and Twitter: Are DMOs successful? *Journal of Destination Marketing & Management*, 16, 100350. <https://doi.org/10.1016/j.jdmm.2019.03.004>
- [36] Leung, X. Y. (2019). Do destination Facebook pages increase fan's visit intention? A longitudinal study. *Journal of Hospitality and Tourism Technology*, 10(2), 2005-2018. <https://doi.org/10.1108/JHTT-06-2018-0041>
- [37] Lin, C. T., Chen, C. W., Wang, S. J., and Lin, C. C. (2018). The influence of impulse buying toward consumer loyalty in online shopping: a regulatory focus theory perspective. *Journal of Ambient Intelligence and Humanized Computing*, 1-11.
- [38] Liu, Y., and Wu, Y. F. (2018). Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).
- [39] Lorch, R. F., and Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology*:

- Learning, Memory, and Cognition*, 16(1), 149. <https://doi.org/10.1037//0278-7393.16.1.149>
- [40] Lowrey, W. (2017). The emergence and development of news fact-checking sites: Institutional logics and population ecology. *Journalism Studies*, 18(3), 376-394.
- [41] Mena, P. (2020). Cleaning up social media: The effect of warning labels on likelihood of sharing false news on Facebook. *Policy & Internet*, 12(2), 165-183. <https://doi.org/10.1002/poi3.214>
- [42] Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., and 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>
- [43] Moravec, P. L., Kim, A., and Dennis, A. R. (2020). Appealing to sense and sensibility: System 1 and system 2 interventions for fake news on social media. *Information Systems Research*, 31(3), 987-1006.
- [44] Moravec, P. L., Kim, A., Dennis, A. R., and Minas, R. K. (2022). Do you really know if it's true? How asking users to rate stories affects belief in fake news on social media. *Information Systems Research*, 33(3), 887-907. <https://doi.org/10.1287/isre.2021.1090>
- [45] Morrow, G., Swire-Thompson, B., Polny, J. M., Kopec, M., and Wihbey, J. P. (2022). The emerging science of content labeling: Contextualizing social media content moderation. *Journal of the Association for Information Science and Technology*, 73(10), 1365-1386. <https://doi.org/10.1002/si.24637>
- [46] Mühlberger, C., Endrejat, P., Möller, J., Herrmann, D., Kauffeld, S., and Jonas, E. (2022). Focus meets motivation: When regulatory focus aligns with approach/avoidance motivation in creative processes. *Frontiers in Psychology*, 13, 4945. <https://doi.org/10.3389/fpsyg.2022.807875>
- [47] Muntinga, D. G., Moorman, M., and Smit, E. G. (2011). Introducing COBRAs: Exploring motivations for brand-related social media use. *International Journal of Advertising*, 30(1), 13-46. <https://doi.org/10.2501/IJA-301-013-046>
- [48] Nassetta, J., and Gross, K. (2020). State media warning labels can counteract the effects of foreign misinformation. *Harvard Kennedy School Misinformation Review*, 1, 1-11. <https://doi.org/10.37016/mr-2020-45>
- [49] Nelson, J. L., and Taneja, H. (2018). The small, disloyal fake news audience: The role of audience availability in fake news consumption. *New Media & Society*, 20(10), 3720-3737. <https://doi.org/10.1177/1461444818758715>
- [50] Nezakati, H., Amidi, A., Jusoh, Y. Y., Moghadas, S., Aziz, Y. A., and Sohrabinezhadtalemi, R. (2015). Review of social media potential on knowledge sharing and collaboration in tourism industry. *Procedia-social and Behavioral Sciences*, 172, 120-125. <https://doi.org/10.1016/j.sbspro.015.01.344>
- [51] Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175-220. <https://doi.org/10.1037/1089-260.2.2.175>
- [52] Pennycook, G., Bear, A., Collins, E. T., and Rand, D. G. (2020a). The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings. *Management Science*, 66(11), 4944-4957. <https://doi.org/10.1287/mnsc.019.3478>
- [53] Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., and Rand, D. G. (2020b). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31(7), 770-780. <https://doi.org/10.1177/095679720939054>
- [54] Pentina, I., Guilloux, V., and Micu, A. (2018). Exploring social media engagement behaviors in the context of luxury brands. *Journal of Advertising*, 47(1), 55-69. <https://doi.org/10.1080/00913367.2017.1405756>
- [55] Petrou, P., Demerouti, E., and Schaufeli, W. B. (2018). Crafting the change: The role of employee job crafting behaviors for successful organizational change. *Journal of Management*, 44(5), 1766-1792. <https://doi.org/10.1177/095679720939054>

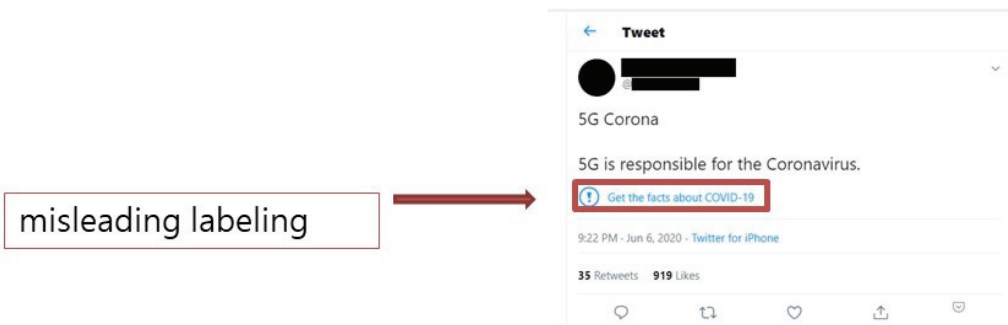
- doi.org/10.1177/0149206315624961
- [56] Salwen, M. B., Garrison, B., and Driscoll, P. D. (Eds.). (2004). *Online news and the public*. Routledge.
- [57] Silverman, C. (2016). This analysis shows how viral fake election news stories outperformed real news on Facebook. *Buzzfeed News*, Retrieved from <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook>
- [58] Solem, B. A. A., and Pedersen, P. E. (2016). The effects of regulatory fit on customer brand engagement: An experimental study of service brand activities in social media. *Journal of Marketing Management*, 32(5-6), 445-468. <https://doi.org/10.1080/0267257X.2016.1145723>
- [59] Solvoll, M. K., and Larrison, A. O. (2020). The (non)use of likes, comments and shares of news in local online newspapers. *Newspaper Research Journal*, 41(2), 204-217. <https://doi.org/10.1177/0739532920919826>
- [60] Song, J., and Qu, H. (2019). How does consumer regulatory focus impact perceived value and consumption emotions? *International Journal of Contemporary Hospitality Management*, 31(1), 285-308. <https://doi.org/10.1108/IJCHM-03-2017-0136>
- [61] Sparks, B. A., Perkins, H. E., and Buckley, R. (2013). Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management*, 39, 1-9. <https://doi.org/10.1016/j.tourman.2013.03.007>
- [62] Swant, M. (2017). Facebook now makes 84% of its advertising revenue from mobile. Adweek. <https://www.adweek.com/performance-marketing/facebook-now-makes-84-of-its-advertising-revenue-from-mobile/>
- [63] Taiwar, S., Dhir, A., Kaur, P., Zafar, N., and Alrasheedy, M. (2019). Why do people share fake news? Associations between the dark side of social media use and fake news sharing behavior. *Journal of Retailing and Consumer Services*, 51, 72-82. <https://doi.org/10.1016/j.jretconser.2019.05.026>
- [64] Toreini, P., Langner, M., Maedche, A., Morana, S., and Vogel, T. (2022). Designing attentive information dashboards. *Journal of the Association for Information Systems*, 23(2), 521-552. <https://doi.org/10.17705/1jais.00732>
- [65] Tretter, S., and Diefenbach, S. (2021). The influence of regulatory focus on media choice in interpersonal conflicts. *Psych*, 3(1), 1-17. <https://doi.org/10.3390/psych3010001>
- [66] Wang, Y., Ma, F., Jin, Z., Yuan, Y., Xun, G., Jha, K., Su, L., and Gao, J. (2018). EANN: Event adversarial neural networks for multi-modal fake news detection. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 849-857).
- [67] Wintersieck, A. L. (2017). Debating the truth: The impact of fact-checking during electoral debates. *American Politics Research*, 45(2), 304-331.
- [68] Xu, H., Teo, H. H., Tan, B. C. Y., and Agarwal, R. (2010). The role of push-pull technology in privacy calculus: The case of location-based services. *Journal of Management Information Systems*, 26(3), 137-176. <https://doi.org/10.2753/MIS0742-1222260305>
- [69] Zaidi, N., Ahuja, V., and Khan, M. N. (2020). Effect of customer characteristics, content characteristics and source credibility on customer response: Developing a conceptual model for the tourism industry. *Journal of Travel & Tourism Research*, 16, 1-24.
- [70] Zhou, Z., Jin, X., Vogel, D. R., Fang, Y., and Chen, X. (2011). Individual motivations and demographic differences in social virtual world uses: An exploratory investigation in Second Life. *International Journal of Information Management*, 31, 261-271. <https://doi.org/10.1016/j.ijinfomgt.2010.07.007>

<Appendix A> YouTube vs. Twitter Practice



authoritative labeling

A1. YouTube Practice: YouTube adds information panels that alert viewers if the source is authoritative (such as when the video comes from an accredited hospital)



misleading labeling

A2. Twitter Practice: Twitter labels content that is misleading or out of context.

<Appendix B>

Items to measure *Believability* with a 7-point scale (adapted from Kim and Denis, 2019)

1. How believable do you find this article?
(1=very little, 4=neutral, 7=very much)
2. How truthful do you find this article?
(1=very little, 4=neutral, 7=very much)
3. How credible do you find this article?
(1=very little, 4=neutral, 7=very much)

Items to measure *Engagement* with a 7-point scale (adapted from Kim and Denis, 2019)

1. How likely are you to 'like' this article?
(1=very little, 4=neutral, 7=very much)
2. How likely are you to leave a comment on this article?
(1=very little, 4=neutral, 7=very much)
3. How likely are you to share this article?
(1=very little, 4=neutral, 7=very much)

Items to measure *Regulatory Focus* with a 5-point scale (adapted from Higgins et al., 2001)

1. Compared to most people, are you typically unable to get what you want out of life?
(1=never or seldom, 3=sometimes, 5=very often)
2. Growing up, would you ever "cross the line" by doing things that your parents would not tolerate?
(1=never or seldom, 3=sometimes, 5=very often)
3. Did you get on your parents' nerves often when you were growing up?
(1=never or seldom, 3=sometimes, 5=very often)
4. Growing up, did you ever act in ways that your parents thought were objectionable?
(1=never or seldom, 3=sometimes, 5=very often)

◆ About the Authors ◆



Hui-Ying Han

Hui-Ying Han is currently employed as a manager at LG Household & Health Care. She holds a Master's degree in Business Administration from the School of Business at Yonsei University. Her research interests include social media and e-commerce.



Youngsok Bang

Youngsok Bang is an associate professor at the School of Business, Yonsei University. Before joining Yonsei University, he was on the faculty of The Chinese University of Hong Kong (CUHK) Business School. His current interests lie in the business value of data analytics, especially in e-commerce, healthcare, fintech, and Internet privacy. His research has appeared in *Asia Pacific Journal of Information Systems*, *Communications of the Association for Information Systems*, *Decision Support Systems*, *Information & Management*, *Information Systems Research*, *International Journal of Information Management*, *Journal of Management Information Systems*, *Management Information Systems Quarterly*, and several other journals.

Submitted: February 20, 2023; 1st Revision: September 3, 2023; Accepted: October 5, 2023