

“I Can’t Respond to Everyone”: An Empirical Investigation of Social Media Influencers’ Management of Controversies

YU-LING CHOU, Department of Computational Media, University of California, Santa Cruz, United States

PEI-HUA TSAI, Department of Computer Science, National Yang Ming Chiao Tung University, Taiwan

GARY HSIEH, Human Centered Design & Engineering, University of Washington, United States

YUNG-JU CHANG*, Department of Computer Science, National Yang Ming Chiao Tung University, Taiwan

Social Media Influencers (SMIs) are vital assets to social media platforms, attracting substantial audiences and driving engagement. However, their heightened visibility also subjects them to intense public scrutiny, leaving them particularly vulnerable when engaging in controversial behavior or making contentious statements. In such moments, SMIs must decide how to respond in ways intended to de-escalate controversy. This study examines how SMIs make strategic decisions following controversial events, the challenges they encounter, and how audiences interpret and react to these responses, revealing critical misalignments between the two perspectives. Through interviews with eight SMIs and fourteen viewer members in a Mandarin-speaking East Asian context, we outline key considerations SMIs weigh during controversies, the significant challenges they encounter, the aspects where audiences commonly express dissatisfaction with SMI responses. Our findings pinpoint three critical areas of misalignment: the interpretation of delayed responses, the selection of target audiences, and the awareness of underlying reasons behind SMIs’ emotional self-protection strategies.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**.

Additional Key Words and Phrases: Social-Media Influencer; Repair strategy

ACM Reference Format:

Yu-Ling Chou, Pei-Hua Tsai, Gary Hsieh, and Yung-Ju Chang. 2026. “I Can’t Respond to Everyone”: An Empirical Investigation of Social Media Influencers’ Management of Controversies. *Proc. ACM Hum.-Comput. Interact.* 10, 6, Article CSCW100 (October 2026), 27 pages. <https://doi.org/10.1145/3816948>

1 Introduction

Social media platforms, burgeoning with digital traffic, have evolved into crucial forums for individuals to share expertise and exert influence. Within this digital landscape, a distinct category of content creators has emerged, often referred to as social media influencers (SMIs) [25, 29, 36, 59]. SMIs build their social capital primarily through self-produced content on digital platforms [28, 32, 37, 50]. This online presence enables them to garner attention, a substantial follower base, and, consequently, financial rewards through sponsorships, advertisements, and collaborations [8, 27]. Simultaneously,

*Corresponding author.

Authors’ Contact Information: Yu-Ling Chou, yhou14@ucsc.edu, Department of Computational Media, University of California, Santa Cruz, Santa Clara, California, United States; Pei-Hua Tsai, phtsai.cs11@nycu.edu.tw, Department of Computer Science, National Yang Ming Chiao Tung University, Hsinchu, Taiwan; Gary Hsieh, garyhs@uw.edu, Human Centered Design & Engineering, University of Washington, Seattle, Washington, United States; Yung-Ju Chang, armuro@nycu.edu.tw, Department of Computer Science, National Yang Ming Chiao Tung University, Hsinchu, Taiwan.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

© 2026 Copyright held by the owner/author(s).

ACM 2573-0142/2026/10-ARTCSCW100

<https://doi.org/10.1145/3816948>

social media platforms rely heavily on these SMIs to create content that attracts people, enhancing user engagement and furthering the platform's profitability.

Nevertheless, like other public figures, SMIs occasionally engage in actions or make statements that may be perceived as inappropriate, immoral, or transgressive [17, 21, 38, 58]. Whether intentional or unintentional, such actions or statements often become the focal point of public scrutiny, giving rise to controversies. In this context, a controversy is defined as a period of heightened public discourse and dispute, characterized by strongly divergent opinions concerning an SMI's actions or statements. We refer to the specific actions or statements that trigger such controversies as controversial actions or statements. These controversies frequently draw substantial attention and, in some cases, escalate into widespread public backlash, with negative sentiments spreading rapidly across social media platforms [14, 64]. Faced with this pressure, SMIs often feel compelled to act swiftly to mitigate the damage and repair their image and reputation. However, the pressure to respond quickly and decisively can sometimes lead to hasty and ill-considered responses that inadvertently exacerbate the controversy, generating even more negative sentiment [64].

The impact of these controversies can extend far beyond the individual SMI, highlighting a pressing issue that merits the attention of the HCI community. At the individual level, many SMIs' careers and livelihoods are deeply intertwined with their public personas. Misjudging audience expectations or failing to navigate a controversy effectively can lead to lasting reputational damage, a loss of followers, and strained relationships with brands and partners. Furthermore, public outrage can take a significant toll on their mental and emotional well-being [34, 53, 55]. Beyond these personal consequences, the controversies may as well impact the organizations, communities, and individuals that are affiliated with the SMIs. In some cases, controversies have led to the cessation or termination of the SMIs' online presence and content creation [34, 55, 64], disrupting not only the influencers themselves but also their followers and the broader audience on social media platforms.

Addressing these challenges is crucial, as SMIs play a pivotal role in shaping online discourse, disseminating information, and fostering platform and community engagement, making them valuable assets and key reasons many people engage with social media platforms [15]. While prior research has examined how SMIs cope with harassment and online attacks from others [9, 31, 34, 35, 45, 53, 55], far less is known about how they manage controversies triggered by their own actions or statements, and how these responses may diverge from audience expectations. This gap motivates closer examination of how socio-technical systems might mediate misalignments between SMIs and their audiences during such moments. Although social media platforms do not necessarily cause these controversies, their affordances for visibility, amplification, and audience interaction shape how such conflicts unfold and persist. As platforms also rely on SMIs as key contributors to content and engagement, understanding how controversies are navigated has direct implications for how platforms might better support creator–audience interaction and sustain a healthy and resilient social media ecosystem.

However, achieving this goal necessitates a deeper understanding of several key areas: how SMIs decide whether and how to respond to controversies, the challenges they face in managing these situations effectively, the critical areas where SMI audience—individuals who occasionally or regularly view SMIs' content—find unsatisfactory, express dissatisfaction with SMIs' responses; and the factors that shape such perceptions in these contexts. By exploring these dimensions and comparing the perspectives of SMIs and their audiences, we aim to uncover potential reasons why some SMIs struggle to resolve controversies, leading instead to their escalation. Understanding these dynamics could help identify the key issues that prevent SMIs from aligning with audience expectations and suggest ways in which system support might alleviate these challenges. To contribute to this knowledge, our research investigates the following research questions:

- RQ1: How do SMIs make decisions regarding whether and how to respond to their controversies?
- RQ2: What challenges do SMIs encounter while navigating controversies?
- RQ3: Which aspects of SMIs' responses to controversies are perceived as most unsatisfactory by their audiences?
- RQ4: What factors shape and influence audience perceptions of SMIs' responses to controversies?

To investigate these research questions, we conducted an in-depth interview study with a total of 22 participants in a Mandarin-speaking East Asian context. This included 8 SMIs who self-reported experiencing controversies related to their behaviors or statements and had taken steps to address them, and 14 SMI audience members who occasionally or regularly consume SMI content and self-reported witnessing multiple SMI's controversies. With the SMI participants, we delved into their decision-making processes, experiences, and the challenges they encountered while navigating the controversies. The interviews with SMI viewers focused on how they perceived, interpreted, and evaluated the SMIs' responses.

The key findings and contributions of this paper are as follows:

- We provide insights into SMIs' decision-making processes during controversies, highlighting the key considerations that shape their response strategies.
- We identify the key aspects that audiences emphasize when assessing SMIs' responses to controversies, as well as the factors that influence these assessments.
- We uncover the challenges that SMIs face in managing their own controversies, and based on these findings, we offer design suggestions for social media platforms to help SMIs better navigate such situations.

2 Related work

2.1 Interaction between SMIs and audience

In recent years, the interaction between online content creators and their audiences has garnered significant scholarly attention. One line of research has focused on content creators' audience management practices [7, 60, 63], examining how creators engage with different types of audiences and manage interactions across multiple platforms [30, 42]. For instance, Wohn and Freeman [63] identified several audience categories conceptualized by streamers, such as family/community, fans, social groups, content-specific viewers, trolls, and lurkers. Their findings reveal that some streamers balance their engagement efforts between general and individual audiences, while others prioritize subscribers or align their strategies with personal interests. In digital gifting, Wang et al. [60] found streamers create exclusive groups for dedicated fans and frequently respond to them. Studies also explore multi-platform engagement strategies [30, 42], with Ma et al. [42] examining cross-platform audience management and Hair et al. [30] highlighting promotional communication with financial supporters.

Complementing these studies, another line of research examines the challenges posed by unpleasant online interactions, including online attacks, harassment, and toxic comments, as well as how content creators, streamers, and SMIs cope with these adversities [9, 31, 34, 35, 45, 53, 55]. For instance, Thomas et al. [53] conducted a survey with 135 creators, finding that online hate and harassment are not only widespread but, for many creators, a regular occurrence. The study highlighted several coping strategies employed by creators, including ignoring harassment, using platform-provided moderation tools, manually reviewing comments, temporarily or permanently leaving the platform, and censoring content to avoid further negativity.

Further research has investigated specific forms of online hate and harassment, as well as the unique challenges faced by targeted groups. For instance, Cai et al. [9] studied "hate raids," a coordinated and malicious activity designed to overwhelm streamers with a sudden influx of hateful messages on platforms like Twitch and Reddit, and proposed platform-level features to mitigate these attacks. Similarly, Meisner [45] interviewed Twitch streamers who experienced hate raids, exploring the tools they utilized and the audience management strategies they employed to combat such attacks.

Additionally, research has highlighted the specific vulnerabilities faced by content creators from marginalized groups, who often encounter targeted forms of harassment and discrimination. For example, Harris et al. [31] examined the racism and discrimination experienced by Black content creators. Uttarapong [55] explored the challenges encountered by Twitch streamers from underrepresented groups, including women and LGBTQ creators, and the strategies they used to navigate and mitigate these adversities. Similarly, Heung et al. [34] combined surveys and interviews to investigate ableist hate and harassment targeting disabled content creators on social media platforms.

To address online harassment and negativity, content creators adopt a range of strategies. Some creators choose to ignore hate speech and avoid engaging with hostile audience members [53], while others attempt to address negativity and harassment, either indirectly or directly. Indirect means often involve using platform moderation tools, such as deleting, blocking, and reporting abusive users [9, 34, 53, 55]. One common technical approach is the use of comment filters—keyword-based systems that automatically screen for offensive language [35]. Recent advances in large language models (LLMs) have significantly expanded the potential of automated moderation, enabling more accurate analysis and detection of hate speech [44, 51, 56]. However, these moderation techniques are not without cost. The opacity of platform moderation and algorithmic governance can raise concerns about potential visibility loss or platform sanctions [23]. As a result, creators may adopt self-censorship as a precautionary strategy—not only to reduce the risk of system-level moderation, but also to proactively avoid content that could trigger negative audience reactions, harassment, or controversy [23, 53].

Beyond technical approach, human moderators are often employed to manage and mitigate such issues [10–12, 41, 55, 62]. Direct responses include clarifying misunderstandings about themselves [34], producing content that counters negative narratives [34], or publicly calling out harassment—often during live streams or on other platforms—to rally support and raise awareness [55].

Even with these proactive measures, the psychological toll of such negative experiences often leads creators to seek emotional support from close ones or professional therapy to help them cope with the psychological impact [53, 55]. In some cases, creators may choose to reduce their visibility by stepping back from social media temporarily [34, 55] to protect their mental health.

Although prior research has examined how SMIs cope with harassment and online hate instigated by others, less is known about how they manage controversies stemming from their own actions or statements. These self-initiated controversies differ in important ways: rather than being framed as external attacks, they require SMIs to confront public criticism directed at their own behaviors—often accompanied by calls for accountability and repair. As a result, responding to such controversies involves a more complex and nuanced approach, where SMIs must weigh personal values, audience expectations, and professional stakes under public scrutiny. This study adopts a dual-perspective approach to address this gap, examining both how SMIs experience and respond to controversies and how audiences interpret and evaluate these responses. Through this lens, we reveal the strategies SMIs employ, the challenges they face, and the reasons some controversies escalate rather than resolve—insights that can inform the design of systems to better support SMIs in navigating such high-stakes situations.

2.2 Public Relation Crisis, Forgiveness, and Image Repair Strategies

Public relations research has examined how controversies—such as transgressions, immoral behavior, and misconduct—can damage the reputations of public figures and organizations, commonly framing these events as public relations crises [4]. Scholars have also studied the effectiveness of various image repair strategies used to restore reputation across contexts [4]. However, most empirical work has focused on organizations and traditional celebrities, with limited attention to SMIs. As a result, SMIs' experiences and perspectives in navigating such crises remain underexplored. To provide context for our research focus, the following sections will briefly review prior work on audience forgiveness, image repair strategies, and recent empirical studies of SMI controversies.

2.2.1 Forgiveness Toward Public Figures and Organizations Involved in Controversies. Scholars have extensively examined how consumers and audiences perceive organizations, brands, and celebrities, particularly focusing on forgiveness and the processing of negative information following their controversial incidents, such as mistakes, misconduct, transgressions, and immoral behavior [16]. A key factor in shaping these perceptions is the nature of audience' and consumers' relationships, commitment, and identification with the organizations and celebrities. For instance, Finsterwalder et al. [26] found that fans' parasocial relationships with celebrities lead to forgiveness behaviors similar to those observed in personal relationships. Both Nikhashemi et al. [46] and Um et al. [54] found that consumers who strongly identify with a brand are more likely to forgive the brand following a public relations crisis. Similarly, Wang and Kim [61] observed that fans with strong identification with a celebrity tend to rationalize and dissociate the celebrity's negative actions more than those with weaker fan identification. Ahluwalia et al. [1] also highlighted that committed consumers often challenge negative information and emphasize positive aspects, whereas those with lower commitment prioritize negative information.

Beyond the relationship aspect, other studies have shown that perceptions of an organization's or celebrity's existing reputation and history also play a crucial role. A history of similar events or a negative prior reputation often suggests a pattern of poor behavior [18]. For example, Rifon et al. [49] found that a celebrity's history of transgressions significantly negatively impacts consumers' attitudes. On the other hand, Dean [22] noted that while organizations with positive reputations may face higher expectations and increased scrutiny, those with negative reputations might receive less severe judgment for similar actions.

2.2.2 Image Repair Strategies. Researchers in public relations have long strived to investigate a range of strategies that organizations and public figures have adopted when facing public relations crises, studying their effectiveness in various situations. Two extensively adopted and acknowledged frameworks are Benoit's Image Repair Theory [3, 4] and Coombs's Situational Crisis Communication Theory (SCCT) [19, 20]. Both frameworks propose a typology of strategies that organizations use to repair their reputations during public relations crises. Common strategies include, but are not limited to, *denial*, *evasion of responsibility*, *reducing offensiveness*, *corrective action*, *mortification or apology*, *bolstering* (offsetting negativity with positive information), and *compensation*. Substantial research has been undertaken to understand these strategies' effectiveness and the factors that influence the public's reception of and reaction to these responses. For example, strategies such as denial, bolstering, and choosing not to respond have generally been deemed less effective [5, 24, 52]. In contrast, approaches like corrective action, mortification, and offering compensation are considered more effective [5, 20]. Additionally, responses that provide explanations of the issues surrounding the crisis tend to be evaluated more favorably than those that do not [20].

Previous studies have also highlighted the influence of various factors on strategy effectiveness. For example, Zhou and Whitla [65] found that whether a controversy is perceived as beyond the

entity's control influences the extent to which the entity's moral reputation is damaged. The impact of crisis severity has also been investigated and plays a role, though to varying extents, depending on who the victims are in the event (consumer vs. company) [19, 39, 40, 47, 66].

2.2.3 Social Media Influencers' Controversies. Prior research has examined SMIs' behaviors that harm their reputations with followers, audiences, or the public. For example, Cocker et al. [17] identified five types of transgressions in the YouTube beauty community: underhand endorsement, over-endorsement, over-emphasis, over-saturation, and over-indulgence. Čop and Culiberg [21] found four actions disliked by Generations Z and Y, including lack of disclosure, change in expertise, reality distortion, and problematic product promotion. Similarly, von Mettenheim and Wiedmann [58] grouped eleven transgressions into three categories: misinformation, hate speech, and privacy violations. A recent study by Ayari et al. [2] examined why users choose to unfollow SMIs and identified twelve recurring reasons, which they grouped into content-related and influencer-related categories. These reasons include, but are not limited to, concerns about content quality, ethical issues, disclosing false information, misalignment in political beliefs or values, offensive or inappropriate behavior, and loss of clear positioning or identity.

However, while these studies provide useful categorizations of controversies and sources of audience dissatisfaction, they primarily focus on what audiences dislike. In contrast, they offer limited insight into how SMIs respond to such situations and how audiences interpret these responses—gaps that our study seeks to address.

To the best of our knowledge, the most relevant study related to our investigation is Wood's analysis [64]. In that study, Woods documented how both influencers faced intensified public backlash after employing response strategies that were perceived as inadequate. Specifically, both were criticized for disabling comment sections, while one was further scrutinized for delaying her response, deleting offensive tweets without explanation, and displaying emotional reactions that some audiences viewed as inauthentic. While the study provides valuable observations of audience reactions and public discourse, it did not examine the controversies from the SMIs' own perspectives. As a result, key questions remain unanswered regarding how these individuals made their decisions, what challenges they encountered, and why they responded in the ways they did. Our study addresses this gap by providing these insights.

3 Methodology

To address our research questions, we conducted in-depth interviews with both SMIs and their audiences—individuals who occasionally or regularly consume SMIs' content. We employed Charmaz's constructivist grounded theory approach [13], which enabled us not only to explore SMIs' experiences in navigating controversies and audiences' perceptions of their responses but also to gain deeper insights into SMIs' decision-making processes, the challenges they face, and the factors influencing audience perceptions of their responses. The following sections detail our methodology. Our study received approval from the institutional review board (IRB), and informed consent was obtained from all participants prior to each interview, ensuring their voluntary participation.

3.1 Recruitment and Participants

3.1.1 Recruitment for SMI Participants. Our recruitment strategy for SMI participants involved two primary approaches: recruiting via a non-profit organization known for its association with content creators and bloggers, and directly contacting SMIs who met our recruitment criteria. Across both approaches, we sought participants who had experienced controversies arising from their behavior or statements within the past two years, sparking online discussions on forums and social media platforms. Initially, we aimed to leverage the non-profit organization as a starting point as well

as a primary channel for identifying eligible participants. However, this approach yielded limited success, resulting in the recruitment of only one participant. Consequently, we adapted our strategy and began reaching out to potential participants directly.

To identify suitable SMIs for the study, our research team conducted a systematic search for posts and news related to SMI controversies on major forums, bulletin board systems, and news outlets within our country. We developed an automated crawler to continuously monitor public online spaces for emerging SMI controversies. The crawler conducted keyword-based searches using general influencer-related terms (e.g., influencer, streamer, YouTuber) and specific creators' channel names, collecting data from multiple sources, including Google News and the bulletin board systems (BBS) with news gossip and influencer-focused boards. When candidate articles were identified, we used a large language model (GPT) to assess whether the event constituted an SMI controversy. The prompt instructed the model to adopt an audience member's perspective and determine whether the SMI's public image would be perceived as positively affected, negatively affected, or unaffected, along with a brief explanation. The crawler ran hourly to enable timely detection, and all screened cases and explanations were reviewed by the research team prior to participant recruitment. Through this process, we compiled a list of potential candidates and sent email invitations to inquire about their willingness to participate in our research.

Ultimately, eight SMIs agreed to participate in the interviews. These participants varied both in the nature and severity of the controversies they had encountered, as well as in their follower counts. Their follower ranges were distributed as follows: two had 10,000–100,000 followers, four had 100,000–500,000, one had 500,000–1 million, and one had over 1 million. This diversity in controversy types and follower ranges enriched our study by offering a broad spectrum of experiences and strategies employed by SMIs in navigating different types of controversies. All of our SMI participants were Mandarin speakers from East Asia, and their content was in Mandarin. Table 1 provides detailed background information on their social media experience, controversies navigated, and platform usage. To ensure the confidentiality and protection of the identities of these SMI participants, we have deliberately omitted specific details that could potentially lead to their identification.

3.1.2 Recruitment for Audience Participants. We recruited audience participants by advertising on Facebook and Instagram, and through the authors' personal networks. Eligible participants were required to occasionally or regularly view SMIs' content and have witnessed multiple SMI controversies. Among those who met these criteria, we prioritized recruiting a diverse sample with variations in demographic characteristics, social media usage patterns, the nature of the controversies they had observed, and the frequency and manner of their interactions with the SMIs involved in these controversies. In total, we recruited 14 participants: 8 aged 20–29, 3 aged 30–39, 2 aged 40–49, and 1 over 50. These participants, from various occupations, were regular consumers of SMI-generated content (see Table 2).

3.2 Interview Process

Our semi-structured interview was originally designed for 120 minutes and conducted via Jitsi Meet¹. Several SMI participants volunteered to contribute more, extending the interview durations up to three to four hours. The interview protocol is provided in Appendix A.

In the audience interviews, we focused on discussing controversial events involving familiar SMIs. Our questions centered on which aspects of SMI responses to controversies were seen as

¹Jitsi Meet is a free, open-source multiplatform voice, video-conferencing and instant-messaging application: <https://meet.jit.si/>.

Table 1. The demographic background of our SMI participants

ID	Age & Gen	SM Exp. ¹	Contr. Count	Contr. Type(s) ²	Followers	Platforms ³
S1	30-34, F	>3y	3-5	(a) (f)	100K-500K	FB, IG, PC, YT
S2	45-49, F	>3y	>5	(a)	10K-100K	BB, FB, IG, LINE, TT/DY, YT
S3	25-29, F	>3y	3-5	(a) (f)	100K-500K	FB, IG, YT
S4	30-34, M	>3y	>5	(c) (d) (e) (f)	1M+	FB, IG, YT
S5	30-34, F	<6mo	3-5	(b)	10K-100K	FB, IG, LINE, YT
S6	30-34, F	>3y	1-2	(b)	100K-500K	FB, IG, PC, TT/DY, YT, Threads
S7	30-34, M	>3y	>5	(d) (e)	500K-1M	YT
S8	- , F	>3y	1-2	(d)	100K-500K	FB, IG, TT/DY, YT

¹ SM Exp.: Social Media Management Experience, Controversies: Number of Controversies, Followers: Follower Size Range, Platforms: The social media platforms managed by SMIs.

² Controversy Type, based on [2]'s proposed typology for unfollowing SMIs: (a) Content quality, (b) Disclosing false information, (c) Ethical concerns, (d) Offensive or inappropriate behavior, (e) Political beliefs misalignment, (f) Values incompatibility

³ BB: Bilibili, FB: Facebook, IG: Instagram, PC: Podcast, TT/DY: TikTok/Douyin, YT: Youtube

⁴ S8 does not disclose her age

⁵ The current table of SMI participants is for review purposes only. To protect their identities from being recognized, we will remove the detailed information about the SMIs in the final, camera-ready version

Table 2. The demographic background of our audience participants

ID	Age & Gen	Occupation	Interaction with SMI	Platform Usage ¹
A1	25-29, M	Financial	One-way Receiving	FB, IG, Ptt, TG
A2	25-29, M	IT	One-way Receiving	FB, IG, YT
A3	40-44, F	Education	Comment, Private Message	DR, FB, IG, Ptt, PC
A4	25-29, F	Medical	One-way Receiving	FB, IG, LINE, YT, NL
A5	50+, F	Unemployed	One-way Receiving	PC, YT
A6	35-39, M	RA	Comment, Live Interaction	BB, DR, FB, IG, Ptt, LINE, WC, YT
A7	35-39, M	Govt EMPL	Comment	DC, FB, IG, Ptt, LINE, TR, TT/DY, WC, YT
A8	40-44, M	Svcs	Comment, Private Message	FB, IG, Ptt, PC, TG, YT
A9	20-24, F	Mfg	One-way Receiving	DR, IG
A10	25-29, F	Govt EMPL	Comment, Like	DR, FB, IG, YT
A11	25-29, M	Student	Comment, Live Interaction	FB, Ptt, YT
A12	20-24, F	Govt EMPL	Comment, Private Message	DR, FB, IG, PC, YT
A13	35-39, M	Student	Comment, Private Message, Live Interaction	DR, FB, IG, Ptt, PC, LINE, YT
A14	25-29, M	Art	One-way Receiving	DR, FB, IG, YT

¹ Platform Usage refers to the platforms where audience participants interact with SMIs. (BB: Bilibili, DC: Discord, DR: Dcard, FB: Facebook, IG: Instagram, NL: Newsletter, PC: Podcast, TG: Telegram, TR: Twitter, TT/DY: TikTok/Douyin, WC: WeChat, YT: Youtube)

most unsatisfactory and what constituted an ideal response. We encouraged participants to reflect on multiple events, comparing their assessments and expectations to reveal what they valued and emphasized in an SMI's response.

The interviews with SMI participants began with a broad discussion about controversies and response strategies observed in other SMIs. When addressing their own controversies, we explored the nature of the events, how they became aware of them, their assessment of the situations, and the response strategies they chose. We also delved into the decision-making processes, including the rationale behind each choice, challenges encountered, and desired platform features. For participants with multiple controversies, we encouraged comparisons to prompt reflection on differences between the events.

For their time and contributions, each participant received compensation of 500 NTD (approximately 16 USD). 4 SMI participants volunteer to donate the compensation.

3.3 Data Collection and Analysis

In alignment with the grounded theory approach, data transcription and analysis were performed concurrently with the data collection phase [13]. Immediately after the initial interviews, we began transcribing and performing open coding. This initial phase of analysis yielded a preliminary set of codes that illuminated gaps in our data. This insight enabled us to refine our recruitment strategy, focusing on participants whose experiences might challenge or further validate our emerging codes and concepts. For instance, our initial set of participants revealed a limitation in their direct interactions with SMIs. Subsequently, we recruited additional participants who had reported more engaged interactions with SMIs. After completing the second round of audience interviews, we shifted our focus to collecting data from SMIs. We diversified the SMI participants by channel topics and follower counts, assuming that these differences might yield varied considerations and strategies. Following the SMI interviews, we initiated a third round of audience recruitment. This final round targeted individuals who could provide deeper insights into specific topics related to the response strategies and concepts discussed by the SMIs, such as the assumption that certain strategies are well-received by audiences. This final stage was designed to enrich our understanding of the dynamics and differences between SMIs' and audiences' perspectives.

All the coding activities took place on Atlas.ti², an online collaborative qualitative-analysis tool. The development of the codebook was an iterative process. Initial code lists were generated from open coding of the first five study sessions, accompanied by memos detailing the codes' and categories' meanings. Researchers then applied these codes independently to the transcripts while continuing to collect data. Throughout this coding process, the team continuously compared and contrasted data, codes, and memos. Regular team discussions led to the iterative revision and refinement of codes and categories as needed. These discussions also help our team to determine the direction for further sampling and to evaluate whether the codes and categories were sufficiently developed and saturated, aligning with the principles of theoretical sampling [13]. This iterative process continued until all team members reached consensus on the final categories and codes. The finalized codebook includes the following categories: audience assessments of SMIs' responses; factors influencing these assessments; audience actions following SMIs' controversies; SMIs' perceptions and evaluations of controversies; SMIs' emotional responses; SMIs' actions; considerations influencing SMIs' decisions; factors shaping SMIs' response strategies; and challenges encountered by SMIs in managing controversies.

In the following findings section, we present the results of our study in response to our research questions. These findings detail SMIs' decision-making processes in addressing controversies, the challenges they encountered, and the audiences' perceptions and evaluations of these response strategies, including the factors that influenced their assessments. A framework encapsulating these findings is illustrated in Figure 1. To present the findings clearly, we reference quotes from our 14 audience participants as A1 through A14, and from our SMI participants as S1 through S8.

4 SMIs' Decision-Making in Responding to Controversies and Challenges Encountered

Our SMI participants employed a variety of strategies in response to the controversies they faced, including remaining silent, blocking comments or disabling the entire comment section, halting their online presence, responding to individual comments, demonstrating improvement, and issuing standalone official statements, most of which align with those reported in the literature [19, 20]. In this section, we present how these participants made decisions regarding their specific responses to controversies and the challenges they encountered while navigating these events.

²Atlas.ti is a computer program used in qualitative research or qualitative data analysis: <https://atlasti.com>.

4.1 SMIs' Decision-Making in Responding to Controversies

After becoming aware of a controversy arising from their behavior or statements, our SMI participants navigated a series of assessments and decision-making processes that influenced and shaped how they responded to the controversies. Below, we detail these processes.

4.1.1 Assessing the Direction of Online Sentiment. The direction of online sentiment played a critical role in shaping whether and how SMI participants responded to emerging controversies. When controversies first arose and overall sentiment was unclear, most participants adopted a "wait-and-see" approach, delaying responses until sentiment trends became more discernible or until they felt confident that a response would be received favorably. Otherwise, they chose not to respond. As one participant explained, *"If you respond, in some ways, you're just fueling the discussion. Especially if you're a big influencer, anything you say will only draw more attention to the issue. So if you really want the situation to calm down, not responding at all might be the better option."* For similar reasons, many participants chose to remain silent when they perceived the overall sentiment to be favorable, aiming to avoid the risk of shifting the conversation in a negative direction.

Several SMI participants reported feeling increased pressure to respond as negative sentiment escalated and attracted broader public attention. While some chose to respond, others continued to remain silent, believing that engagement would be unproductive. Participants cited either a lack of compelling arguments in their defense or the perception that public opinion was too entrenched to change. As S6 noted, *"People will only believe what they want to believe."* This perception led to concerns that responding could not only fail to shift sentiment but might worsen the situation. S4 expressed frustration, saying, *"Anything I say would be magnified and attacked relentlessly."*

For participants who chose to respond, the presence of sufficient positive sentiment was critical. Endorsements from experts or professionals increased their confidence to take a more assertive stance. Support was assessed not only through comments but also through engagement metrics such as likes, views, shares, and subscriptions. Several participants noted that even when comments appeared predominantly negative, positive sentiment was reflected in these metrics. As S1 said, *"Those who agree with me don't necessarily comment, but they express their support through likes, subscriptions, and shares."* Seeing an increase in these metrics made them feel that there was a "silent majority," giving them the sense that public opinion was not as unfavorable as it appeared.

4.1.2 Key Purposes for Responding to Controversies and Tailoring Response Strategies. SMI participants articulated a range of purposes for responding to controversies. These purposes shaped both the content of their responses and the strategies they adopted. We describe these purposes and the associated response strategies in the sections that follow.

Mitigating Controversy to Protect Professional and Personal Relationships: A common motivation for responding was to mitigate potential harm to affiliated parties, such as employers, sponsors, or collaborators. Participants expressed concern that leaving a controversy unaddressed could escalate and result in professional consequences or reputational damage for those they worked with. S3, for instance, stated, *"I probably can't take a controversial stand. It would greatly affect the reputation of the company I work for."* Similarly, S7, mindful of his team's future, shared, *"I feel it's my duty to think about everyone's future. I need to expand and grow because these people help me achieve what I want, and thus I have a responsibility and obligation to fulfill their dreams."*

Participants in this category described a strong sense of obligation to demonstrate accountability, even when they did not feel personally at fault. For them, apologizing was less about conceding

specific criticisms and more about acknowledging responsibility for having triggered the controversy. As S6 explained, *"Sincerity is I apologizing to [the victim] and letting everyone see I take [the victim] seriously. Whether the victim forgives me or not isn't something I particularly think about, I just have to do what I am expected to do."* S3 similarly stated, *"My apology was primarily because I felt I had wasted societal resources and caused my audience to doubt their trust in me."* S8 echoed this sentiment, adding, *"Sometimes, an apology doesn't necessarily mean you're in the wrong, but rather it's a display of your larger capacity for understanding and tolerance. It's about taking a step back for the sake of the bigger picture."*

Safeguarding a Carefully Constructed Public Persona: For many participants, protecting a carefully constructed public persona was a central priority, as it was seen as essential to maintaining their fanbase and overall online persona. When controversies arose, these SMIs viewed their fans and long-term followers as the primary audience to address. As S5 explained, *"You have to distinguish between important and unimportant [audiences]. [...] What the unimportant ones say doesn't concern me. I have to pay more attention in handling the matters that are from my fans."* S7 echoed this sentiment, noting, *"I don't want to disappoint my fans and audience who have supported me for a long time."*

However, the decision to respond was also influenced by the SMIs' perceptions of what their persona required. In some cases, participants chose not to respond to minor criticisms, as they believed their audience expected them to *"focus on more significant issues"* (P7). This alignment with their online persona also extended to the style and tone of their response. S8, for example, adopted a soft, apologetic tone—even when feeling unfairly criticized—to preserve her gentle image: *"My persona is not about being harsh or brutally honest [...] My channel does not accept negativity."*

Defending Perceived Truths in the Face of Public Criticism: For some SMIs, the reason to respond stemmed from a desire to defend what they believed to be the truth, rather than to calm the controversy or offer an apology. These participants often resisted public pressure to apologize, feeling that their stance was justified. As S1 put it, *"Some people messaged me suggesting I change my statement to avoid criticism. I refused because I believe I am right [...] What I want to express is that the truth remains the truth. I cannot apologize for the truth."* For these SMIs, public disapproval did not equate to wrongdoing; they believed that apologies were only warranted when they were genuinely at fault. S6 remarked, *"Apologize only if you're wrong, and don't if you're not. You need to be clear-headed and not blindly follow public opinion."*

In other cases, responses were motivated by a need to correct what participants perceived as false accusations. For example, S5 chose to speak out when accused of fabricating content, *"I get more upset if someone accuses me of faking things than being greedy. Honestly, yes, I'm greedy. Who isn't? But the accusation of faking really hurt me because it's a complete distortion of the truth."*

Leveraging Controversy for Increased Attention and Visibility: Not all SMI participants viewed controversies as inherently negative. Some treated them as opportunities to increase visibility, strategically using responses—or silence—to amplify attention. S1 remarked, *"Stirring things up is a good thing. The real problem is failing to do so [...] The worst-case scenario is following the standard apology process and getting no response."* S5, other other hand, used silence strategically to sustain public curiosity, explaining, *"I chose not to respond. [...] Whether the accusations are true or not, let them speculate. The more they speculate about me, the more attention I receive."*

Seeking Emotional Support and Sympathy: Lastly, a notable reason for some SMI participants in responding to controversies was the desire to seek sympathy and emotional support from their

fans and followers. Several participants adopted an emotional tone in their responses, aiming to foster a sense of personal connection with their audience. As S3 explained, *“Expressing sadness can lead to more personal and emotional discussions, which is quite effective. It can trigger fans to defend you in the comments.”* By tapping into their followers’ emotions, these participants hoped to cultivate loyalty and create a supportive environment during times of public scrutiny.

In some cases, SMI participants went beyond seeking sympathy—they actively encouraged their fans to disseminate their message. S8 described this as a strategic move to amplify their voice, noting, *“It’s really about letting [the fans] know you’re suffering [...] They could spread your words, which can be really helpful.”* These responses, designed to generate empathy and solidarity, allowed SMIs to build a stronger, more emotionally engaged fanbase, which they believed was crucial for navigating controversies.

4.1.3 Deciding the Platform to Offer the Response. Many SMI participants carefully considered which platform to use when responding to a controversy. This decision was shaped by where the controversy first gained traction, perceptions of audience receptiveness across platforms, and platform-specific affordances. While some participants preferred to respond on the platform where the issue originated, others chose platforms with more supportive audiences to avoid, as S4 described, *“engaging in battles outside their territory.”* For instance, S3 explained, *“I feel the audience on Instagram interacts more closely with me [...] That’s where you can see more defending-type of fans.”*

SMIs also strategically considered platform affordances. S6, though mainly active on YouTube, opted for quicker text-based posts on Facebook and Instagram: *“Writing a post was quicker than making a video. Plus, this didn’t require a video.”* In contrast, S5 chose YouTube for its emotional expressiveness, *“Emotions, expressions, even a sigh, can subtly enhance an apology.”* Others selected platforms based on functionality. S3 preferred responses without lasting traces, *“A post stays forever [...] Once it’s over, it should be over.”*

Several SMI participants who managed personas across multiple platforms deliberated whether to differentiate platform roles or present a unified response. For example, S4 favored differentiation, explaining, *“My YouTube channel is for discussing [the material]. No matter how significant the controversy, I almost never create videos to address it.”* Conversely, S5 adopted a more unified approach, stating, *“It’s crucial to present a synchronized image to ensure the audience perceives what’s actually happening. If my Instagram continues to show me enjoying coffee and fine dining, people might question whether I’m truly affected by a certain event.”*

4.2 Challenges Faced by SMI Participants in Managing Controversies

When navigating controversies, SMI participants encountered several challenges that constrained timely and effective decision-making. Our analysis identified four key challenges: delayed awareness of emerging controversies, difficulty distinguishing constructive feedback from irrelevant or harmful criticism, the emotional toll of sustained criticism and attacks, and challenges in deriving actionable insights from platform metrics.

4.2.1 Delayed Awareness of an Emerging Controversies. A pervasive initial challenge was delayed recognition that participants’ actions or statements had become controversial. Many SMIs became aware of controversies only after they had already attracted significant public attention, often extending beyond their typical audience. This late awareness was described by S3, who noted, *“When friends outside of this field start showing concern, asking ‘are you okay?’. That’s when I realized things have become serious.”* S6 experienced a similar situation, describing it as being in *“a parallel universe,”* where she noted, *“People around me, including my audience, had already seen it. [...] They assumed I was silently handling it, appearing strong and unaffected, but in reality, I was simply unaware of*

it.” This lag in awareness often prevented timely responses, leading to situations where public opinion had already escalated markedly. S3 stated, *“By the time I noticed the event, it already had thousands of responses [...] The earlier comments tend to get more likes and rise to the top. Essentially, top comments largely dictate the direction of public opinion. [...] The public sentiment has already been determined, and it’s too late to reverse it.”*

4.2.2 Distinguishing Feedback Deserving Attention. SMI participants encountered significant challenges in discerning which feedback warranted attention after a controversy. A major difficulty was navigating the overwhelming volume of irrelevant or toxic comments, often from casual viewers or “clueless onlookers” who had no prior engagement with their content. Amid this noise, participants aimed to prioritize feedback from their core audience—those who regularly followed and interacted with them—and sought out constructive criticism, even from non-followers, if it was expressed respectfully. Yet locating such valuable feedback amidst the influx of hostile or irrelevant comments was a major challenge. Supportive or constructive responses were often buried in a sea of negativity, requiring participants to sift through blame and off-topic attacks—an effort that was both time-consuming and mentally draining.

S7 expressed his frustration when facing this challenge, stating, *“Every time something like this happens, it attracts a crowd of onlookers. They’re like sharks; they’re not here to determine whether this thing is right or wrong. They don’t care if this thing is real or fake, or whether you’re right or wrong. They don’t care if you have a point. They are just here to watch the show, and for these people, anything you do or say is wrong. They will endlessly criticize.”* S5 echoed this sentiment and expressed a desire for platform assistance, noting, *“The process [of discerning fans and onlookers] is very tedious, so I hope the platform would be willing to assist me in this.”*

The issue of sockpuppet accounts further complicates the situation. These accounts are frequently viewed as tools for manipulating conversations or propagating particular narratives, thereby posing a challenge for SMI participants in discerning authentic engagement.

4.2.3 Emotional Toll of Managing Overwhelming Criticism and Backlash. Nearly all participants highlighted the emotional burden imposed by such backlash. Although a few participants mentioned becoming more resilient over time, learning to dismiss certain attacks, they still vividly recalled the distress and hurt they experienced during their earlier, less experienced stages. Common emotional struggles included frustration, self-doubt, and deep distress. S8 shared, *“When everyone started criticizing, I began to question if I was really that bad, wondering how such a simple thing could be perceived as fake. It made me doubt whether my skills were as poor as they thought.”* For some, the emotional toll was so profound that it led them to consider pausing or even abandoning their online presence altogether. They questioned whether continuing to engage with their audience was worth the effort, given the intense scrutiny and the ongoing challenge of maintaining their well-being. S7 stated, *“I started wondering whether what I’ve been doing for so long is right or wrong. I even began to think that maybe those [audience] who left were right, and perhaps those who stayed are mistaken.”*

This struggle was especially difficult for participants who felt that, while their actions or statements may have been inappropriate, the intensity of the backlash was disproportionate. They found it challenging to reconcile the harshness of public criticism with their unintentional mistakes, which further amplified their emotional distress.

4.2.4 Navigating the Disconnection Between Metrics and Audience Sentiment. SMI participants noted a disconnect between surface-level metrics like likes and dislikes and the actual sentiments of their audience. These metrics lacked context, making it difficult to interpret negative reactions—especially when they conflicted with comment content. S5 explained, *“If someone clicks ‘dislike’ and then comments, I wish I could easily see, ‘Oh, they disliked and commented.’ [...] I want to see what the*

majority of dislikes are about—was it the pace of the video, or was the content politically incorrect? That way, I can focus on optimizing those aspects in my future content.”

With an understanding of how SMIs navigate controversies and the challenges they face, it is crucial to now shift our focus to the audience’s perspective. Exploring how audiences perceive and react to SMI responses helps us identify potential misalignments and uncover the factors that contribute to audience dissatisfaction.

5 Audience’ Perception and Assessment of SMI’s Response Strategies

In this section, we address the research questions focused on the audience’s perspective, exploring which aspects of SMIs’ responses the audience deems crucial for effectively addressing controversies, as well as the factors that shape and influence these perceptions.

5.1 Seven Critical Aspects in SMI Controversy Response from the Audience’s Perspective

Our analysis identified seven key aspects that audiences evaluate when assessing the effectiveness of SMIs’ responses to controversies. These aspects are critical in shaping the audience’s perception of whether the response is sufficient and appropriate—and are often sources of audience dissatisfaction. We detail each aspect below.

5.1.1 Targeting Responses at the Appropriate Audience. Audience participants emphasized that responses should be directed to the appropriate audience. Many found it reasonable for SMIs to prioritize core followers during controversies to preserve trust within their primary audience. For example, A9 noted, *“Many people support and like you. For the part where you let them down, those people deserve an apology.”* However, when controversial actions directly affected specific individuals, participants expected SMIs to address those parties first. Prioritizing a general audience in such cases was viewed as inappropriate or insincere. As A8 critiqued one SMI’s approach, *“He should have first sought forgiveness from those directly affected, rather than us. [...] Resolve the issues you’ve caused before crying to us about it.”* If the response appeared to target only regular followers or fans, participants perceived it as avoiding accountability. As A8 put it, SMIs were *“just comforting each other in their own bubble.”*

5.1.2 Addressing Core Issues in Responses. Audience participants frequently expressed criticism towards responses that did not sufficiently address the central issues of the controversy. A common observation was that some SMIs tended to offer generalized apologies, which appeared disconnected from the specific concerns raised. These apologies often gave the impression that SMIs were apologizing out of a sense of obligation, rather than genuinely addressing the underlying problems. As A9 remarked, *“It feels like he doesn’t really know what upset everyone, just trying to quell the storm.”* Similarly, A12 noted, *“He knew people were upset, but not why, and then he apologized for something he had no idea about.”* This perceived lack of specificity often led to audience disappointment and disengagement. There was a broad consensus among participants that SMIs should explicitly list and address critical issues in their responses to demonstrate awareness, understanding, and a genuine willingness to resolve each issue.

When SMIs defended their actions, audience participants expected concrete evidence or clear justification. As A9 emphasized, *“Let everyone see the evidence. Use evidence to let people know that you were falsely accused.”* A5 explained, *“It’s not about apologizing; it’s about whether people accept your explanation.”*

5.1.3 Maintaining Transparency Throughout the Response. Audience participants consistently emphasized the importance of transparency and clarity in SMI responses, particularly in explaining the context and circumstances of a controversy. Many expressed frustration when explanations were vague or delayed. As A10 noted, *"I think they should quickly explain the situation clearly and concisely after it happened."* Because audiences often learned about controversies through secondary sources (e.g., news or social media discussions), a lack of firsthand explanation left them unable to assess the situation, increasing expectations for transparent communication from the SMI.

Participants also expected transparency to be sustained over time. Responses that were difficult to access or removed quickly were perceived lacking transparency. For instance, A9 criticized temporary apologies, saying, *"An apology statement that's only up for a short time is like not having one at all. The next day it's gone, so it's pointless."* Such practices were commonly interpreted as insincere and negatively shaped audience perceptions of the response.

5.1.4 Expressing Sufficient Mortification in Responses. Audience participants further emphasized that when SMIs were perceived to have acted inappropriately or unethically, they were expected to explicitly acknowledge wrongdoing and express genuine remorse. Failure to do so often raised doubts about whether SMIs recognized their mistakes or would change their behavior. A9 noted: *"If everyone thinks you're wrong, and there's evidence to prove it, but you still insist on your stance, whether they are fans or just normal audience, they will gradually lose interest and stop watching your videos."*

5.1.5 Demonstrating Correction and Improvement. When SMIs engaged in inappropriate, transgressive, or immoral behavior, audience participants sometimes also expected clear corrective actions and concrete plans to prevent recurrence. As A9 suggested, *"Describe the points you need to apologize for, and then tell us how you won't make these mistakes again."* Many participants expressed skepticism toward verbal commitments alone, emphasizing that apologies without action were insufficient. A5 noted, *"If there's no real action, no matter how many times you apologize, people who aren't convinced won't change their minds."*

Expectations for corrective action varied by context. In cases where misinformation was disseminated, participants expected prompt and accurate rectification. A14 emphasized, *"Update immediately on what's correct. Then, explain the content of the update."* In situations involving potential plagiarism, participants anticipated specific remedial measures. A9 elaborated, *"They should turn off the video's monetization since it's someone else's hard work. [...] And [the SMI] should edit their video description to reference the original source, letting people know it's not entirely original."*

5.1.6 Displaying Authenticity and Prudence in Response Manner. Audience participants emphasized authenticity and prudence as essential qualities of effective responses. Many preferred video or livestream formats, which were seen as better conveying sincerity and emotional nuance. A4 noted, *"It allows for eye contact and communication. Text alone can be easily misunderstood."* However, some participants cautioned against dramatic or emotional displays, perceiving them as manipulative. A11, for example, critiqued such responses as *"like dramatizing for grabbing attention,"* while A9 warned, *"[Excessive emotion] might just show that the emotions are unprocessed, or seeking sympathy, which could have the opposite effect."*

Additionally, careful crafting of a response was deemed essential, signaling the SMI's seriousness towards the issue. Instances of poor language usage in responses led to criticisms about the SMI's attitude, as A9 pointed out, *"This made me question his seriousness and whether he was just being careless."* A prevalent suggestion among participants was to have responses reviewed by others to preempt potential backlash.

The underlying motives behind responses were also closely scrutinized, as they are key to signaling authenticity. Delays in issuing responses were commonly interpreted as indicative of ulterior motives, with A5 associating such delays with a tendency to “*downplay the controversy*,” and A4 remarking, “*A delayed apology can feel like waiting for public anger to subside or for the issue to be forgotten.*”

5.1.7 Willingness to Receive and Consider Feedback. Lastly, audience participants strongly criticized behaviors that signaled reluctance to receive or consider feedback, such as deleting comments or disabling comment sections. These actions were widely perceived as evasive and non-communicative. As A10 noted, “*It’s very negative when they refuse to face the situation or shut down all possible ways for people to question and challenge them.*” P11 also observed that audiences quickly adapt to these tactics, stating, “*People will just go elsewhere to comment. They might leave comments on your previous videos or on other YouTube videos.*” Participants strongly emphasized the importance of SMIs being open to feedback to demonstrate their receptiveness and willingness to understand the issues at hand.

5.2 Factors Influencing Audience Participants’ Perception and Assessment

We identified several key factors that influenced the audience’s evaluations of SMI responses, outlined below.

5.2.1 Perceived Nature of the Controversy. Participants’ expectations of SMI responses were strongly shaped by how they perceived the nature and severity of the controversy. They commonly distinguished between minor, correctable mistakes and more serious transgressions. For less severe issues—such as unintentional misinformation or controversial opinions without deceptive intent—participants generally held lower expectations. As A4 explained, “*It’s impossible to always have perfect information.*” In such cases, audiences primarily expected acknowledgment and correction, as A9 stated, “*A simple post acknowledging the mistake, explaining that it was an oversight, that’s all that’s needed.*”

However, when controversies involved actions perceived as immoral or illegal, particularly those undermining an SMI’s integrity, expectations were substantially higher. Participants anticipated both legal accountability and a sincere public apology, and were markedly less forgiving, as A4 said, “*It changes how you see that person. Even if the SMI apologizes or whatever, you know that’s their true nature. So, no matter how they apologize, I don’t think it really changes anything.*”

5.2.2 Audiences’ Pre-Existing Perception of the SMI’s Image. Audience participants’ pre-existing perceptions of an SMI strongly shaped how they evaluated both controversies and responses. SMIs known for being provocative or frequently involved in controversies were often held to different—and sometimes lower—standards. Illustrating this, A9 commented, “*People are accustomed to him being blunt and using harsh language on his channel. So, if he were to suddenly apologize, it might seem odd to his audience, like ‘Why is he doing this now?’*” A13 noted a similar sentiment about a humor-oriented SMI: “*His way of responding was quite humorous, and he addressed it in a similar manner. [...] It’s in line with the black humor or the witty style he usually employs in his YouTube videos.*”

However, when a controversy or a response contradicted the pre-existing image of an SMI, participants often experienced confusion, concern, or disappointment. This sense of disappointment could escalate to feelings of ‘deception’ and ‘betrayal,’ especially when the contradicted image was a key reason for their following the SMI and when participants held a high aspirational image of the SMI. A11, for example, expressed this disappointment on one SMI’s plagiarism, “*You begin to doubt their actual talent. And if they’re not as skilled as you believed, what’s the point of watching*

their YouTube? It feels like a betrayal." A1 shared his letdown regarding an SMI he held in high regard: *"People like me have thought of that [SMI] as an intelligent and wise person. So, when he responded in this way, I felt it's not up to the level we expect from him."*

Audience perception of an SMI's image often extends to interpreting their intentions behind managing controversies and crafting responses. A7, for example, speculated on an SMI's response timing, *"The timing fits with what I've observed about his channel's usual approach [...] He may think that simultaneously issuing a response statement and making new content could affect his editing schedule, 'why not just combine the new content and response into one video?' This approach is advantageous for him in terms of marketing and time efficiency."* Similarly, A3 speculated on an SMI's apparent indifference to criticism, *"He didn't really care to respond to the criticism [...] I think his logic is, 'It doesn't matter. Let them criticize; having traffic first is more important.'"* However, repeated controversies and patterned responses led to audience fatigue. A9 expressed this sentiment, *"Too many controversies, so many that it feels like 'another controversy, another apology.' It makes people feel tired and not wanting to watch anymore."*

5.2.3 Perceived Influence of SMI. There was a general consensus that with greater influence comes greater responsibility, as P14 commented, *"He's become a public figure. So, in turn, he has this responsibility to not just go around shooting off his laser gun. [...] The voice of a public figure is a powerful tool. I hope he handles it well, not turning [the debate] into unnecessary fights."* They were especially concerned about the impact on impressionable segments of the audience who might lack the capacity to critically evaluate the SMI's information. For example, A9 highlighted this concern, *"There are people who really can't judge that information properly. They'll think, 'Hey, if you say it's good, it must be. Let's give it a try.' And that could lead to health issues for them."* Extending this concern to younger audiences, P9 noted, *"The content is watched by young audiences, including elementary and high school students. It's important to let them know this is not okay to emulate the behavior."*

5.2.4 Audiences' Commitment and Identification with the SMI. Finally, we observed that the level of commitment and identification audience participants had with SMIs played a role in shaping some audience participants' responses to the SMIs' incidents and their subsequent responses. Among our study participants, those who identified themselves as followers or fans displayed a tendency to rationalize the SMIs' behavior, often dissociating any misconduct from the content they enjoy. This led to a continued positive perception of the SMI, despite any immoral actions they might have engaged in. A4's perspective on an SMI accused of plagiarism illustrates this point, *"I didn't really care much [about the SMI's event] [...] I think I don't care maybe because I just want to hear him talking. I enjoy their storytelling and interaction. It's not that I support plagiarism. But, but I don't know why, [his controversy] just didn't bother me much."*

6 Discussion

While prior research has extensively investigated how SMIs manage online hate, harassment, and attacks instigated by others [9, 31, 34, 35, 45, 53, 55], much less is known about how they handle controversies triggered by their own actions or statements. Our study extends this line of work by focusing on self-initiated controversies, in which SMIs themselves become the source of public backlash. Unlike unprovoked hate or harassment, these controversies arise from behaviors perceived as inappropriate, unethical, or transgressive, requiring SMIs to actively engage with audience feedback to address the underlying criticism and repair their reputations.

Although prior work suggests that some SMIs engage in self-censorship to reduce negativity [53], our findings show that such strategies are often insufficient. Many self-initiated controversies were unintentional, with SMIs perceiving the content as acceptable at the time, and only recognizing its

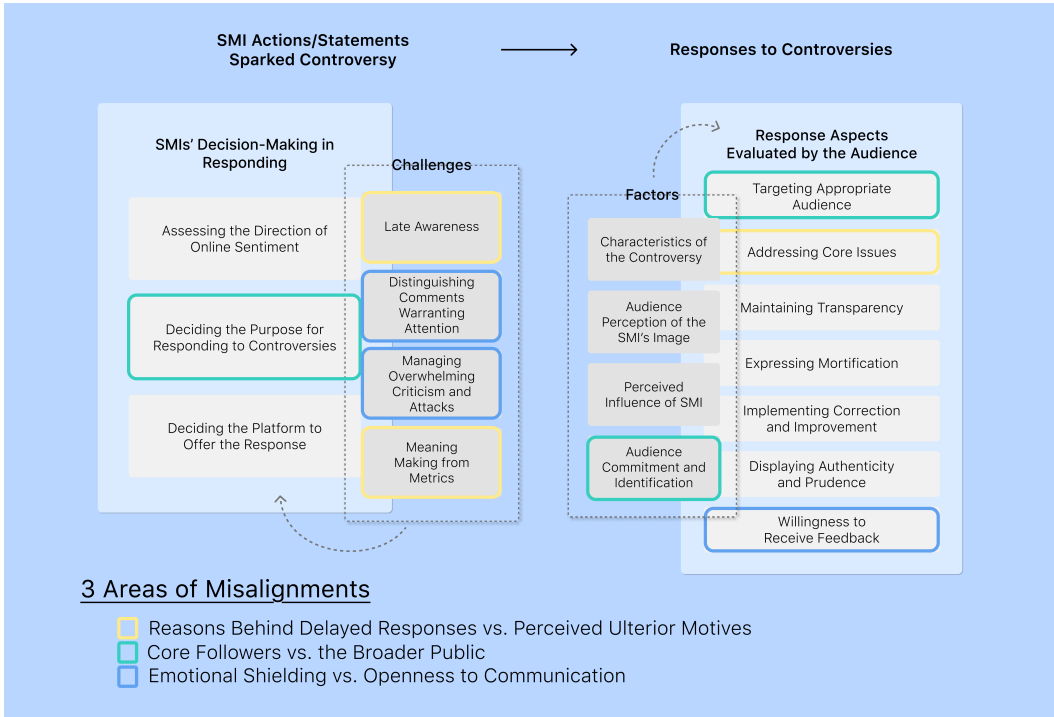


Fig. 1. This framework summarizes the insights for managing controversies from both the perspectives of SMIs and their audiences. It highlights three key aspects of SMIs' decision-making processes when responding to controversies, as well as the challenges they encounter. Additionally, the figure presents the audiences' perceptions and evaluations of SMIs' response strategies, along with the factors that influence these assessments. The figure also illustrates three areas of misalignment between SMIs' decision-making processes and audience expectations, showing how SMIs' internal decision-making and challenges, together with audience evaluation criteria and contextual factors, give rise to these misalignments.

controversial nature after it had attracted broader public attention. Once surfaced, negativity often spread beyond core followers, amplifying its scale and visibility. As a result, SMIs are compelled to treat these incidents as public relations crises [4], necessitating response strategies distinct from those used to manage routine online hate or harassment.

By incorporating the perspectives of both SMIs and their audiences, our study offers new insights into the nuanced decision-making processes, considerations, and challenges SMIs face when managing self-initiated controversies. Our findings help explain why some controversies persist or escalate, contributing to a deeper understanding of SMI-audience dynamics in these contexts. While some alignment emerged between the two perspectives, we identified three key areas of misalignment that appeared to drive escalation. These stemmed from misunderstandings, with audience participants often interpreting SMIs' motivations in ways that were not intended, and SMI participants often overestimating the effectiveness or reception of their responses. We present these areas of misalignment and illustrate them in Figure 1.

6.1 Reasons Behind Delayed Responses vs. Perceived Ulterior Motives

A recurring challenge for SMIs lies in how their delayed responses to controversies are interpreted by their audiences. Prior research by Woods [64] has shown that audiences often perceive such delays as

evasive, leading to further backlash. Our interviews with audience participants reinforced this view, with many interpreting delayed responses as signs of insincerity or avoidance. Some speculated that SMIs were deliberately waiting for public outrage to diminish, hoping the controversy would be forgotten over time.

However, our interviews with SMI participants revealed more nuanced reasons for these delays. While a few SMIs admitted to postponing their responses intentionally, the majority cited practical challenges rather than deliberate avoidance. A key factor mentioned was delayed awareness of the controversy itself. Often, by the time SMIs became aware of the issue, it had already escalated significantly. This was particularly common when controversies originated outside their primary platforms, such as on community-driven forums or third-party social media sites, with SMIs relying on personal networks to alert them to brewing issues. As a result, delays were often the result of limited awareness rather than intentional inaction.

Even within their own platforms, SMIs faced the challenge of monitoring vast amounts of activity across multiple channels. Managing interactions, tracking metrics, and interpreting sentiment were difficult. Without clear insights into the reasons behind certain metrics, such as spikes in dislikes or negative comments, SMIs sometimes dismissed these as routine social media dissent, a common occurrence in the social media environment [31, 34, 53, 55], rather than recognizing them as early warning signs of a controversy.

These challenges were further compounded by SMIs' desire to assess the overall direction of public sentiment and the potential impact of a response, leading many to adopt a "wait-and-see" approach. Although participants recognized the need to address the situation, they hesitated due to uncertainty about how to respond effectively and concerns that an ill-considered response might escalate the controversy or fail to mitigate negative sentiment. However, this approach may unintentionally reinforce the audience's belief that the delay is motivated by ulterior motives. Importantly, our findings suggest that SMIs may not need to fear that any response will necessarily worsen the situation. Audience participants consistently emphasized that they valued seeing the core issue addressed. Therefore, rather than hesitating out of concern that a response might aggravate the controversy, SMIs could benefit from showing a willingness to directly engage with the issue. By doing so, they may help prevent further audience skepticism regarding their motives or sincerity. Taken together, these findings highlight the need to better support SMIs in recognizing and interpreting emerging controversies. We discuss related design implications in a later section.

6.2 Core Followers vs. the Broader Public

The second misalignment concerns SMIs' tendency to prioritize the expectations of their committed and identity-aligned core followers over those of a broader, more diverse public. Prior research has explored how SMIs manage relationships with their dedicated audiences, emphasizing the emotional connections and reciprocal support between SMIs and their core followers [63]. Our study extends this work by revealing that, in the context of controversies, SMIs face significant challenges in adapting their audience management strategies. This difficulty arises because the audience for their responses to controversies is no longer limited to their original followers and viewers but often includes the broader public. Addressing this new and diverse group necessitates a shift in strategy, particularly in managing the tension between satisfying the expectations of their core followers and those of the broader public—a tension that has not been examined in prior research.

Specifically, our findings highlight that while SMIs often prioritize addressing their core followers, assuming that these highly committed audiences, due to their familiarity with the SMI's persona and values, can better contextualize both the controversy and the response, this focus can unintentionally exacerbate public discontent. As controversies attract wider attention, a broader audience—typically

lower in commitment and lacking the same level of familiarity or loyalty—becomes involved and is more likely to provide candid and critical feedback on the key issues requiring attention. In contrast, loyal followers, as highlighted in prior research and observed in our interviews, may be more inclined to rationalize or overlook the SMI's negative information [1, 61], potentially obscuring the core issues driving the controversy. While supportive, this dynamic may hinder SMIs from quickly identifying and addressing the underlying concerns contributing to public discontent.

Thus, if SMIs focus solely on feedback from their loyal followers—who often provide positive reinforcement—they risk missing valuable opportunities to de-escalate the situation more effectively. Our findings suggest that this narrow focus can be perceived by the wider audience as an attempt to remain within an echo chamber, rather than genuinely engaging with the controversy. Such perceptions may leave key issues unresolved and could further alienate the broader public, potentially intensifying the controversy instead of diffusing it. Taken together, our findings suggest that effectively navigating widely publicized controversies requires SMIs to move beyond a narrow reliance on their core followers and engage with a broader range of audience perspectives. However, doing so is inherently challenging, as SMIs must balance public accountability with maintaining a persona that resonates with their loyal audience. These tensions point to the need for better support in helping SMIs understand and respond to diverse audiences. We elaborate this later in the design-implication section.

6.3 Emotional Shielding vs. Openness to Communication

The third notable misalignment arises from the tension between SMIs' need to shield themselves emotionally and their audience's expectation for openness in addressing criticism. To protect their mental well-being during controversies, many SMIs adopted defensive strategies—such as seeking emotional support from loyal followers or temporarily disabling comments to avoid negative feedback.

While emotionally shielding strategies may offer short-term relief, they risk undermining the very openness that audiences value [15]. Particularly when controversies stemmed from SMIs' own actions or statements, audience participants viewed emotional appeals as deflecting responsibility and interpreted comment moderation or deletion as a refusal to engage with legitimate criticism rather than an act of self-protection [15]. As prior work has shown [64], and as our findings corroborate, perceived withdrawal from engagement rarely de-escalates conflict; instead, it shifts critique to other platforms, where it may intensify and become less manageable.

This interpretive divide aligns with broader debates surrounding the “cancel culture” [57] and public accountability. As research has shown [57] notes, many audiences regard public criticism as a necessary mechanism for holding influential figures accountable, while those targeted often experience such responses as disproportionate punishment or unjust social sanction. Within this context, SMI participants' attempts at emotional shielding—understood by them as self-protection—can be reinterpreted by audience as resistance to accountability, legitimizing calls for punitive actions such as unfollowing, boycotts, or public denunciation, intensifying the cycle of conflict.

At the same time, maintaining openness during controversies is inherently challenging for SMIs. To read and engage with criticism, SMIs must endure a barrage of personal attacks, which imposes a substantial emotional burden. As a result, SMIs often need time to shield themselves emotionally before they can re-engage with their audiences in a productive and meaningful way. But audiences may not be fully aware of the emotional toll that SMIs experience during controversies or the necessity for SMIs to take time to recover before re-engaging. This lack of mutual understanding contributes to misunderstandings, highlighting the potential for fostering greater empathy from audiences regarding the emotional challenges SMIs face. Such empathy could alleviate pressure

on SMIs during difficult periods, potentially reducing the likelihood of withdrawal from content creation [9, 45, 64], which would benefit both SMIs and their audiences.

Taken together, these findings reveal a fundamental tension in controversy management: while audiences expect openness and accountability, SMIs must also protect their emotional well-being to sustain continued engagement. Navigating this balance is challenging, as efforts to shield oneself may be misinterpreted as avoidance, further intensifying conflict. This suggests the need for support that enables meaningful engagement while reducing unnecessary emotional burden and fostering greater mutual understanding between SMIs and their audiences, as we discuss in the following section.

6.4 Implications for Social Media Platforms

Our findings highlight the need for social media platforms to better support SMIs in navigating controversies by addressing gaps in awareness, feedback interpretation, audience understanding, and emotionally sustainable engagement. Across the three misalignments identified above, platforms are positioned not only to detect and moderate harmful content, but also to help creators and audiences better understand one another during moments of conflict.

First, platforms could improve SMIs' awareness and understanding of potential controversies. The first misalignment showed that delayed responses are often shaped not only by hesitation, but also by limited awareness and difficulty interpreting early signals. Predictive models built from indicators such as increased negativity in comments and sudden spikes in engagement metrics may help identify possible controversies at an earlier stage. Early notifications could improve SMIs' ability to recognize emerging issues and respond while negativity remains manageable. In addition, platforms could provide richer insights into the reasons behind audience reactions. Prior research on streamers' analytic needs [43] highlights the value of qualitative feedback alongside engagement metrics, suggesting an opportunity for LLMs to summarize audience concerns from comments. Platforms might also proactively solicit structured feedback from audience members when abnormal events are detected, allowing more constructive input than what may surface through comments alone.

Second, platforms could better support SMIs in understanding and managing heterogeneous audiences during controversies. As shown in the second misalignment, controversies often draw in new and less-committed audiences whose expectations differ from those of committed followers. Tools that differentiate between core followers and broader publics—based on engagement patterns or prior interaction history—may help SMIs contextualize feedback. Aside from comparing how their comments and engagement behaviors differ from core followers, insights about the newcomers may also be collected from how these broader audiences have reacted in similar prior events on other channels. These insights may help SMIs prioritize issues that are broadly shared while allowing them to organize response by audience types to better address each segment's expectations.

Third, platform interventions should go beyond basic functions like filtering or blocking toxic comments [34, 53, 55], and instead support more nuanced engagement with criticism. The third misalignment suggests that simply shutting out criticism may reduce immediate harm, but can also intensify perceptions of avoidance. Building on recent research that utilizes LLMs to enhance the accuracy of hate speech detection across diverse contexts, events, and cultural nuances [44, 51, 56], we suggest that moderation tools distinguish between uncivil language and substantive criticism that reflects deeper concerns or systemic issues. In addition, platforms might also leverage LLMs to soften the tone of harsh comments or extract the core concerns, helping SMIs engage with feedback without unnecessary emotional strain. However, LLM-based rephrasing must be carefully calibrated—if not, it may distort the original meaning or be perceived as inauthentic, underscoring

the need for transparency and oversight in deployment. Future research should further examine how SMIs and audiences perceive and respond to such moderation strategies.

Finally, platforms could actively guide audiences toward more constructive engagement with SMIs. For example, one of the largest social media platforms, X³, has tested prompts encouraging users to pause and reconsider potentially harmful or offensive replies before posting them. While these prompts can deter impulsive harmful replies, their impact may be limited if users ignore the warnings or become desensitized to such interventions over time. Building on this approach, we propose not just prompting reconsideration but actively steering users toward offering helpful feedback or actionable suggestions. For instance, instead of a generic warning, platforms could present specific examples of constructive responses or offer templates that frame disagreement as problem-solving. Such prompts could be paired with social cues—such as highlighting comments rated as helpful by creators or other users—to reinforce desirable behavior. While such interventions may not fully eliminate prompt fatigue or harmful discourse, incorporating contextual cues and emphasizing positive engagement norms may increase the salience of the prompt and reduce the likelihood of it being ignored. Encouraging such constructive feedback would not only increase the likelihood of SMIs addressing the core issues but also create more meaningful interactions between SMIs and their audience, contributing to a more positive environment for resolving controversies.

Taken together, these implications suggest that platforms can play a more active role in mediating creator–audience dynamics during controversies. Rather than focusing solely on content removal or abuse prevention, platforms may also help creators recognize emerging controversies, understand diverse audiences, and engage more sustainably with criticism. In doing so, they may reduce the misalignments identified in our study and support healthier forms of controversy management online.

7 Study Limitations

This study has several limitations. First, our sample included only SMIs who had experienced potential controversies within the year prior to participation and did not include SMIs with fewer than 10K followers. Many participants also self-identified primarily as educational or informational content creators, which may bias our findings toward the perspectives and strategies of this subgroup. Additionally, self-selection bias [33] may be present, as SMIs willing to participate and share their experiences may differ systematically from those who declined. Responses may also be influenced by social desirability bias [48], as participants may have presented themselves in a favorable light during interviews. Second, our findings are shaped by the cultural specificity of our sample, which consisted exclusively of Mandarin-speaking SMIs and audience members. While this focus enabled in-depth examination within a particular linguistic and cultural context, it may limit generalizability to other settings. Cultural differences—such as contrasts between collectivist and individualistic orientations—may influence how SMIs interpret public image and respond to controversies [6]. Future work should examine more culturally diverse samples to assess the extent to which the misalignments identified here are culturally contingent or generalizable across contexts.

8 Conclusion

This study examined how SMIs navigate self-initiated controversies and the challenges they face in managing public responses. By analyzing SMIs' decision-making processes, we highlight the tensions they encounter in balancing the expectations of core followers and broader audiences. Our findings point to three areas where support is needed: improving awareness of emerging

³X is one of the world's largest social media platforms: <http://x.com>

controversies, helping SMIs identify meaningful audience feedback, and mitigating harmful criticism without obscuring constructive input. We also show the importance of aligning SMIs' response strategies with audience interpretations, while fostering greater audience understanding of the emotional and strategic constraints SMIs face. Together, these insights advance understanding of SMI controversy management and highlight opportunities for platforms to better support healthier creator–audience dynamics.

Acknowledgments

We deeply appreciate the contributions of all participants in our study. This work was funded by the National Science and Technology Council, Taiwan (113-2222-E-A49 -009 -MY3).

References

- [1] Rohini Ahluwalia, Robert E Burnkrant, and H Rao Unnava. 2000. Consumer response to negative publicity: The moderating role of commitment. *Journal of marketing research* 37, 2 (2000), 203–214.
- [2] Salma Ayari, Khoulood Oueslati, and Imène Ben Yahia. 2025. Exploring the causes to unfollow social media influencers: A qualitative study. *Journal of Human Behavior in the Social Environment* 35, 3 (2025), 347–365.
- [3] William L Benoit. 1997. Image repair discourse and crisis communication. *Public relations review* 23, 2 (1997), 177–186.
- [4] William L Benoit. 2024. *Accounts, Excuses, and Apologies: Image Repair Theory Extended*. State University of New York Press.
- [5] William L Benoit and Shirley Drew. 1997. Appropriateness and effectiveness of image repair strategies. *Communication reports* 10, 2 (1997), 153–163.
- [6] Stephen Bochner. 1994. Cross-cultural differences in the self concept: A test of Hofstede's individualism/collectivism distinction. *Journal of cross-cultural psychology* 25, 2 (1994), 273–283.
- [7] Ross Bonifacio, Lee Hair, and Donghee Yvette Wohn. 2023. Beyond fans: The relational labor and communication practices of creators on Patreon. *New Media & Society* 25, 10 (2023), 2684–2703.
- [8] Gillian Brooks, Jenna Drenten, and Mikolaj Jan Piskorski. 2021. Influencer celebrification: How social media influencers acquire celebrity capital. *Journal of Advertising* 50, 5 (2021), 528–547.
- [9] Jie Cai, Sagnik Chowdhury, Hongyang Zhou, and Donghee Yvette Wohn. 2023. Hate raids on twitch: Understanding real-time human-bot coordinated attacks in live streaming communities. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–28.
- [10] Jie Cai and Donghee Yvette Wohn. 2021. After violation but before sanction: Understanding volunteer moderators' profiling processes toward violators in live streaming communities. *Proceedings of the ACM on Human-computer Interaction* 5, CSCW2 (2021), 1–25.
- [11] Jie Cai and Donghee Yvette Wohn. 2022. Coordination and Collaboration: How do Volunteer Moderators Work as a Team in Live Streaming Communities?. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [12] Jie Cai, Donghee Yvette Wohn, and Mashael Almoqbel. 2021. Moderation visibility: Mapping the strategies of volunteer moderators in live streaming micro communities. In *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*. 61–72.
- [13] Kathy Charmaz. 2014. *Constructing grounded theory*. sage.
- [14] Grace Y Choi and Anne Marie Mitchell. 2022. So sorry, now please watch: Identifying image repair strategies, sincerity and forgiveness in YouTubers' apology videos. *Public Relations Review* 48, 4 (2022), 102226.
- [15] Yu-Ling Chou, Hsuan-Jen Lee, Jie Tsai, En-Chi Liang, and Yung-Ju Chang. 2023. I Like Their Autonomy and Closeness to Me: Uncovering the Perceived Appeal of Social-Media Influencers. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [16] Emily Chung and Michael Beverland. 2006. An exploration of consumer forgiveness following marketer transgressions. *ACR North American Advances* (2006).
- [17] Hayley Cocker, Rebecca Mardon, and Kate L Daunt. 2021. Social media influencers and transgressive celebrity endorsement in consumption community contexts. *European Journal of Marketing* 55, 7 (2021), 1841–1872.
- [18] Timothy Coombs. 2013. Situational theory of crisis: Situational crisis communication theory and corporate reputation. *The handbook of communication and corporate reputation* (2013), 262–278.
- [19] W Timothy Coombs. 1998. An analytic framework for crisis situations: Better responses from a better understanding of the situation. *Journal of public relations research* 10, 3 (1998), 177–191.
- [20] W Timothy Coombs. 2007. Protecting organization reputations during a crisis: The development and application of situational crisis communication theory. *Corporate reputation review* 10 (2007), 163–176.

- [21] Nina Grgurić Čop and Barbara Culiberg. 2020. Business is business: The difference in perception of influencer's morality between generation Y and Z. In *Advances in Digital Marketing and eCommerce: First International Conference, 2020*. Springer, 56–61.
- [22] Dwane Hal Dean. 2004. Consumer reaction to negative publicity: Effects of corporate reputation, response, and responsibility for a crisis event. *The Journal of Business Communication* (1973) 41, 2 (2004), 192–211.
- [23] Brooke Erin Duffy and Colten Meisner. 2023. Platform governance at the margins: Social media creators' experiences with algorithmic (in) visibility. *Media, Culture & Society* 45, 2 (2023), 285–304.
- [24] Sujay Dutta and Chris Pullig. 2011. Effectiveness of corporate responses to brand crises: The role of crisis type and response strategies. *Journal of Business Research* 64, 12 (2011), 1281–1287.
- [25] Nadja Enke and Nils S Borchers. 2021. Social media influencers in strategic communication: A conceptual framework for strategic social media influencer communication. In *Social media influencers in strategic communication*. Routledge, 7–23.
- [26] Jörg Finsterwalder, Thomas Yee, and Alastair Tombs. 2017. Would you forgive Kristen Stewart or Tiger Woods or maybe Lance Armstrong? Exploring consumers' forgiveness of celebrities' transgressions. *Journal of Marketing Management* 33, 13-14 (2017), 1204–1229.
- [27] Karen Freberg, Kristin Graham, Karen McGaughey, and Laura A Freberg. 2011. Who are the social media influencers? A study of public perceptions of personality. *Public relations review* 37, 1 (2011), 90–92.
- [28] Sophia Gaenssle and Oliver Budzinski. 2021. Stars in social media: New light through old windows? *Journal of Media Business Studies* 18, 2 (2021), 79–105.
- [29] Jan-Frederik Gräve and Annika Greff. 2018. Good KPI, good influencer? Evaluating success metrics for social media influencers. In *Proceedings of the 9th International Conference on Social Media and Society*. 291–295.
- [30] Lee Hair, Ross Bonifacio, and Donghee Yvette Wohn. 2022. Multi-platform practices among digital patronage creators. *Convergence* 28, 5 (2022), 1438–1456.
- [31] Camille Harris, Amber Gayle Johnson, Sadie Palmer, Diyi Yang, and Amy Bruckman. 2023. "Honestly, I Think TikTok has a Vendetta Against Black Creators": Understanding Black Content Creator Experiences on TikTok. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–31.
- [32] Alison Hearn and Stephanie Schoenhoff. 2015. From celebrity to influencer: Tracing the diffusion of celebrity value across the data stream. *A companion to celebrity* (2015), 194–212.
- [33] James J Heckman. 1990. Selection bias and self-selection. In *Econometrics*. Springer, 201–224.
- [34] Sharon Heung, Lucy Jiang, Shiri Azenkot, and Aditya Vashistha. 2024. "Vulnerable, Victimized, and Objectified": Understanding Ableist Hate and Harassment Experienced by Disabled Content Creators on Social Media. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–19.
- [35] Shagun Jhaver, Quan Ze Chen, Detlef Knauss, and Amy X Zhang. 2022. Designing word filter tools for creator-led comment moderation. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–21.
- [36] S Venus Jin, Aziz Muqaddam, and Ehri Ryu. 2019. Instafamous and social media influencer marketing. *Marketing Intelligence & Planning* 37, 5 (2019), 567–579.
- [37] Susie Khamis, Lawrence Ang, and Raymond Welling. 2017. Self-branding, 'micro-celebrity' and the rise of social media influencers. *Celebrity studies* 8, 2 (2017), 191–208.
- [38] Belinda Kintu and Karim Ben-Slimane. 2020. Companies responses to scandal backlash caused by social media influencers. *International Journal of Market Research* 62, 6 (2020), 666–672.
- [39] Daniel Laufer, Kate Gillespie, Brad McBride, and Silvia Gonzalez. 2005. The role of severity in consumer attributions of blame: Defensive attributions in product-harm crises in Mexico. *Journal of International Consumer Marketing* 17, 2-3 (2005), 33–50.
- [40] Betty Kaman Lee. 2005. Hong Kong consumers' evaluation in an airline crash: A path model analysis. *Journal of Public Relations Research* 17, 4 (2005), 363–391.
- [41] Na Li, Jie Cai, and Donghee Yvette Wohn. 2023. Ignoring as a moderation strategy for volunteer moderators on twitch. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–7.
- [42] Renkai Ma, Xinning Gui, and Yubo Kou. 2023. Multi-platform content creation: the configuration of creator ecology through platform prioritization, content synchronization, and audience management. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [43] Keri Mallari, Spencer Williams, and Gary Hsieh. 2021. Understanding analytics needs of video game streamers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [44] Sarah Masud, Sahajpreet Singh, Viktor Hangya, Alexander Fraser, and Tanmoy Chakraborty. 2024. Hate Personified: Investigating the role of LLMs in content moderation. *arXiv preprint arXiv:2410.02657* (2024).
- [45] Colten Meisner. 2023. Networked responses to networked harassment? Creators' coordinated management of "hate raids" on Twitch. *Social Media+ Society* 9, 2 (2023), 20563051231179696.

- [46] SR Nikhashemi, Rowan Kennedy, and Felix Mavondo. 2023. Leveraging virtual brand community engagement and consumer brand identification as a bounce-back brand recovery strategy: Role of brand endorsers. *Journal of Marketing Communications* (2023), 1–41.
- [47] Sun-A Park and María E Len-Ríos. 2010. Who suffers? The effect of injured party on attributions of crisis responsibility. *The handbook of crisis communication* (2010), 591–606.
- [48] Wendy L Richman, Sara Kiesler, Suzanne Weisband, and Fritz Drasgow. 1999. A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of applied psychology* 84, 5 (1999), 754.
- [49] Nora J Rifon, Mengtian Jiang, and Shuang Wu. 2023. Consumer response to celebrity transgression: investigating the effects of celebrity gender and past transgressive and philanthropic behaviors using real celebrities. *Journal of Product & Brand Management* 32, 4 (2023), 517–529.
- [50] Theresa M Senft. 2013. Microcelebrity and the branded self. *A companion to new media dynamics* (2013), 346–354.
- [51] Xinyue Shen, Yixin Wu, Yiting Qu, Michael Backes, Savvas Zannettou, and Yang Zhang. 2025. HateBench: Benchmarking Hate Speech Detectors on LLM-Generated Content and Hate Campaigns. *arXiv preprint arXiv:2501.16750* (2025).
- [52] Rajesh Kumar Srivastava. 2019. Rebuilding a global brand under crisis—case of a global brand Maggi. *International Journal of Pharmaceutical and Healthcare Marketing* 13, 2 (2019), 118–139.
- [53] Kurt Thomas, Patrick Gage Kelley, Sunny Consolvo, Patrawat Samermit, and Elie Bursztein. 2022. “It’s common and a part of being a content creator”: Understanding How Creators Experience and Cope with Hate and Harassment Online. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–15.
- [54] Nam-Hyun Um. 2013. Celebrity scandal fallout: How attribution style can protect the sponsor. *Psychology & Marketing* 30, 6 (2013), 529–541.
- [55] Jirassaya Uttarapong, Jie Cai, and Donghee Yvette Wohn. 2021. Harassment experiences of women and LGBTQ live streamers and how they handled negativity. In *Proceedings of the 2021 ACM International Conference on Interactive Media Experiences*. 7–19.
- [56] Nishant Vishwamitra, Keyan Guo, Farhan Tajwar Romit, Isabelle Ondracek, Long Cheng, Ziming Zhao, and Hongxin Hu. 2024. Moderating new waves of online hate with chain-of-thought reasoning in large language models. In *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 788–806.
- [57] Emily A. Vogels. 2021. Americans and “cancel culture”: Where some see calls for accountability, others see censorship, punishment. <https://www.pewresearch.org/internet/2021/05/19/americans-and-cancel-culture-where-some-see-calls-for-accountability-others-see-censorship-punishment/>
- [58] Walter von Mettenheim and Klaus-Peter Wiedmann. 2023. Influencer transgressions: The impacts on endorser and brand. *Journal of Media Economics* 35, 1-2 (2023), 28–62.
- [59] Demetris Vrontis, Anna Makrides, Michael Christofi, and Alkis Thrassou. 2021. Social media influencer marketing: A systematic review, integrative framework and future research agenda. *International Journal of Consumer Studies* 45, 4 (2021), 617–644.
- [60] Dennis Wang, Yi-Chieh Lee, and Wai-Tat Fu. 2019. “I love the feeling of being on stage, but I become greedy” Exploring the impact of monetary incentives on live streamers’ social interactions and streaming content. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–24.
- [61] Sai Wang and Ki Joon Kim. 2020. Consumer response to negative celebrity publicity: The effects of moral reasoning strategies and fan identification. *Journal of Product & Brand Management* 29, 1 (2020), 114–123.
- [62] Donghee Yvette Wohn. 2019. Volunteer moderators in twitch micro communities: How they get involved, the roles they play, and the emotional labor they experience. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–13.
- [63] Donghee Yvette Wohn and Guo Freeman. 2020. Audience management practices of live streamers on Twitch. In *Proceedings of the 2020 ACM International Conference on Interactive Media Experiences*. 106–116.
- [64] Chelsea Woods. 2021. You “Can’t Put Concealer on This One”. *Research Perspectives on Social Media Influencers and Their Followers* (2021), 129.
- [65] Lianxi Zhou and Paul Whitla. 2013. How negative celebrity publicity influences consumer attitudes: The mediating role of moral reputation. *Journal of Business Research* 66, 8 (2013), 1013–1020.
- [66] Ziyuan Zhou and Eyun-Jung Ki. 2018. Does severity matter?: An investigation of crisis severity from defensive attribution theory perspective. *Public Relations Review* 44, 4 (2018), 610–618.

A Interview Protocol

This appendix presents the semi-structured interview protocol used in our study. The protocol was used flexibly, with follow-up questions asked as appropriate.

A.1 Interview Structure for Social Media Influencers

A.1.1 *Background and Channel Context.*

- Which platforms do you primarily create content on?
- What type(s) of content do you usually produce on these platforms?
- How would you describe your content style or online persona?
- How do you usually engage with audience feedback, and how does this feedback influence your future content or channel management? (e.g., comments, private messages, live interactions, analytics or backend metrics)

A.1.2 *Experiences with Controversies and Response Strategies.*

- Could you describe any controversies you have experienced in the past? (Please describe what happened and what triggered the controversy.)
- How did you respond when the controversy first emerged?
- What factors influenced your decision to respond in that particular way?
- Have you experienced multiple controversies? If so, how did your response strategies differ across events, and why?
- Did your response strategies differ across platforms? If so, why?
- After the controversy, how did you adjust your content creation or channel management strategies, if at all?
- What motivated these post-controversy changes?
- Overall, how would you describe the impact of this controversy (or these controversies) on you personally and professionally?

A.1.3 *Closing.* Today we discussed your experiences with controversies, your response strategies, and how these experiences influenced your subsequent channel management. Is there anything else you would like to add or elaborate on regarding today's discussion?

A.2 Interview Structure for Audience Members

A.2.1 *Media Consumption and Interaction Habits.*

- What types of influencers or content creators do you usually follow or watch?
- How do you typically interact with influencers? (e.g., watching only, commenting, liking, private messaging, participating in live streams)

A.2.2 *Perceptions of Influencer Controversies and Responses.*

- Can you recall any influencer controversies that left a strong impression on you? (Please describe what happened.)
- What were your thoughts or reactions to these events?
- Did your reactions differ across different types of controversies? If so, why?
- How did you perceive the influencer's response to the controversy?
- What did you think about how they responded initially?
- What did you think about how they managed the situation afterward?
- Across different events, how did your evaluation of influencers' responses differ, and why?
- Besides the event itself, were there other factors that influenced how you evaluated the influencer's response?(e.g., your expectations of the influencer, their public image, or your relationship with them)
- How did you personally respond to changes in the influencer's behavior or content strategy after the controversy?
- What influenced your reaction?

A.2.3 Closing. Today we discussed your perspectives on influencer controversies and how influencers respond to them. Is there anything else you would like to add or share that we have not covered?

Received May 13, 2025; revised January 13, 2026; accepted March 17, 2026