An integrated model of ambivalence and deterrence of bystander inaction in delegitimizing doxing

Li, Yang-Jun; He, Zhi-Zhi; Cheung, Christy M.K.; Shen, Xiao-Liang; Lee, Matthew K.O.

Published in:
Information & Management

Published: 01/04/2025

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

Published version (DOI):
10.1016/j.im.2025.104118

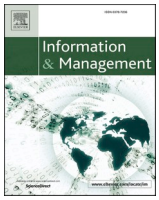
Publication details:
Li, Y.-J., He, Z.-Z., Cheung, C. M. K., Shen, X.-L., & Lee, M. K. O. (2025). An integrated model of ambivalence and deterrence of bystander inaction in delegitimizing doxing. *Information & Management*, *62*(3), Article 104118. https://doi.org/10.1016/j.im.2025.104118

Information & Management 62 (2025) 104118

Contents lists available at ScienceDirect

Information & Management

journal homepage: www.elsevier.com/locate/im



An integrated model of ambivalence and deterrence of bystander inaction in delegitimizing doxing

Yang-Jun Li [a], Zhi-Zhi He [b,*], Christy M.K. Cheung [c], Xiao-Liang Shen [d], Matthew K.O. Lee [b]

[a] School of Management, Beijing Institute of Technology, 5, Zhongguancun South Road, Beijing, 1000081, China
[b] Department of Information Systems, City University of Hong Kong, 83, Tat Chee Avenue, Kowloon Tong, Hong Kong, 999077, China
[c] Department of Management, Marketing and Information Systems, Hong Kong Baptist University, 224, Waterloo Road, Kowloon Tong, Kowloon, Hong Kong, 999077, China
[d] School of Information Management, Wuhan University, 299, Bayi Road, Wuhan, 430072, China

ARTICLE INFO

Keywords:
Doxing
Bystander effect
Attitudinal ambivalence
General deterrence theory
Polynomial regression and response surface analysis
Cyberbullying

ABSTRACT

Our study seeks to explain bystander inaction by guiding the regulation of delegitimizing doxing. We develop a contextualized research model by integrating attitudinal ambivalence theory and general deterrence theory. Our results innovatively uncover three congruence effects describing how positive and negative appraisals are combined to determine bystander inaction. Moreover, perceived SNS deterrence weakens the negative impact of attitudinal univalence on bystander inaction. As one of the pioneering empirical inquiries into delegitimizing doxing, our study contributes to theory and practice related to bystander intervention in delegitimizing doxing.

1. Introduction

A spectre is haunting the Internet—the spectre of cyberbullying. With more people spending time online, cyberbullying, which refers to the use of technology to harass, threaten, embarrass, or target another person [1], is one of the top safety concerns for Internet users in today's digital world. It is reported that over 40 % of adult internet users and 60 % of children in the United States have experienced cyberbullying in their lifetime [2]. Cyberbullying has a devastating impact on victims, resulting in various negative outcomes such as anxiety, self-abasement, depression, concentration difficulty, substance abuse, delinquency, and even suicide [3]. This issue has received widespread social attention, and governments and social networking sites (SNSs) have introduced policies and measures to deter cyberbullying activities [2]. In cyberbullying incidents, bystanders represent the largest group of stakeholders, as compared to perpetrators and victims, with estimates of between 60 % and 70 % of Americans having witnessed cyberbullying [4]. However, although bystanders are considered to have great potential to stop and prevent cyberbullying, only a minority of bystanders intervene in cyberbullying [3,5].

Bystander inaction is generally frowned upon by most bystander intervention researchers and practitioners, as it may exacerbate the severity of cyberbullying incidents [6,7]. Prior studies have extensively examined how to get bystanders to stand up rather than stand by [5,8]. Bystanders who don't respond to cyberbullying have also been criticized for their lack of empathy and responsibility [4,9]. However, the emergence of a unique form of cyberbullying, i.e., delegitimizing doxing, challenges the current understanding of bystander inaction. Delegitimizing doxing describes the act of releasing personal details onto the Internet to shame and humiliate the target, and it often portrays the target as a norm transgressor and provides evidence of moral violations [10]. Given that people are motivated to express outrage on social media to signal their moral quality once they find a norm has been violated [11], delegitimizing doxing is rampant on social media. Examples of delegitimizing doxing include exposing private sexual relationships of celebrities [12], exposing animal abusers [13], and exposing uncivilized behavior in public places [14].

Unlike bystander inaction in general cyberbullying, bystander inaction in delegitimizing doxing is not necessarily undesirable. Delegitimizing doxing posts usually have moralized language and thus go viral easily on social media [15,16]. The resulting widespread dissemination of personal information and moral violations will likely ruin the target's

Conflict of Interest: All authors declare that they have no conflicts of interest.

* Corresponding author.
E-mail address: zhizhihe2-c@my.cityu.edu.hk (Z.-Z. He).

https://doi.org/10.1016/j.im.2025.104118
Received 1 April 2024; Received in revised form 4 February 2025; Accepted 6 February 2025
Available online 7 February 2025

personal, public, and professional life, causing disproportional harm to the target [17]. With no responses to delegitimizing doxing posts, bystander inaction has the potential to cool down the delegitimizing doxing incident because social media platforms algorithmically prefer to push posts with lots of discussions. As such, bystander inaction can shape the magnitude of the harm to the target caused by delegitimizing doxing. Without understanding bystander inaction in delegitimizing doxing, bystanders may be unintentionally encouraged by the content of doxing and current community standards of SNSs to join in and escalate delegitimizing doxing. Thus, it behooves researchers to examine the antecedents of bystander inaction in delegitimizing doxing to guide the development of bystander intervention strategies for delegitimizing doxing.

Against this backdrop, the main research objective of our study is to develop a contextualized research model to explain bystander inaction in delegitimizing doxing. Our research question is: how do the attitudinal ambivalence towards delegitimizing doxing and the SNS deterrence influence bystander inaction in delegitimizing doxing? To answer the research question, we have made the following attempts. First, with problematizing the current understanding of bystander inaction in delegitimizing doxing, we draw upon attitudinal ambivalence theory [18] and argue that bystander behaviors are the result of their experienced attitudinal ambivalence, which is defined in our study as the simultaneous existence of positive and negative appraisals of delegitimizing doxing. The attitudinal ambivalence theory is well suited for this study because it is developed to understand human attitudes and behaviors towards ambiguous norms, multifaceted objects, and hybrid experiences [40,42–44]. Delegitimizing doxing is a double-edged sword in that it holds the alleged transgressors to account and enforces social norms [10, 19], and on the other hand, it inevitably violates the target's privacy and dignity and may cause disproportionately harmful impact on the target [10,20]. Unlike the context where good and evil are clearly defined, bystanders are likely to generate both sympathy for the victim and understanding of the perpetrator in delegitimizing doxing due to the subjectivity of moral judgment. As such, bystanders are very likely to form positive and negative evaluations of delegitimizing doxing at the same time, leaving them not knowing how to respond to delegitimizing doxing. Thus, unlike prior studies that attributed bystander inaction to causes, including diffusion of responsibility, evaluation apprehension, pluralistic ignorance, and lack of empathy [4,21], attitudinal ambivalence theory offers a good starting point to understand the uniqueness of bystander inaction in delegitimizing doxing. This also responds to the calls to leverage the problematization approach to develop our research model [22,23].

Second, by combining the lens of individuals and environment to provide a rich understanding of an online deviant behavior, we apply the general deterrence theory [24] to understand how social media platforms' countermeasures against delegitimizing doxing intervene in the effects of attitudinal ambivalence on bystander inaction in delegitimizing doxing. Deterrence from social networking sites (SNS deterrence) is defined as the degree of severity, certainty, and celerity of platform sanctions perceived by users if violating community standards and policies. Platforms have adopted a series of regulations and enforcement to form their deterrence, such as community standards, algorithms detection, expert review, education center, and built-in reporting function. Due to the complex nature of delegitimizing doxing, how social media platforms regulate it is a challenging task and the enforcement of the countermeasures against delegitimizing doxing is also likely to vary across platforms, forming environmental conditions where bystanders respond to delegitimizing doxing. According to interactionism, individuals' behavior, especially unethical or illegal behavior (e.g., delegitimizing doxing), is the result of the interplay between personal and environmental factors [25]. Thus, attitudinal ambivalence and perceived SNS deterrence will collectively determine bystander inaction in delegitimizing doxing, offering a more integrative explanation of bystander inaction in delegitimizing doxing.

Using an online survey with 542 samples, we leverage polynomial regression and response surface analysis (PRRSA) to provide nuanced understanding of how the combinations of positive and negative appraisals (i.e., attitudinal ambivalence) affect bystander inaction in delegitimizing doxing. Our study uncovers three influencing congruence effects and the moderating effect of SNS deterrence. For theoretical contributions, as one of the pioneering empirical inquiries into delegitimizing doxing, our study responds to the calls [33,75] to differentiate forms of cyberbullying and is expected to stimulate more academic discourses on this significant but underexplored phenomenon (i.e., delegitimizing doxing). Second, by delineating the distinctive antecedents of bystander inaction in delegitimizing doxing from previous studies, our work not only contributes to bystander intervention research by challenging the current understanding of bystander inaction and its causes in cyberbullying literature, but also underscores the value of SNS deterrence to mitigate delegitimizing doxing. Third, our study offers a more rigorous and innovative approach (i.e., PRRSA) to test attitudinal ambivalence theory and specifies the boundary condition within which attitudinal ambivalence determines bystander inaction, thus advancing attitudinal ambivalence research. For practical contributions, our findings offer implications for policymakers and platforms to distinguish delegitimizing doxing from general cyberbullying and underscore the proactive role platforms can play by implementing robust detection, clear countermeasures, and strong enforcement to mitigate delegitimizing doxing.

In the following sections, we introduce the conceptualization of delegitimizing doxing, bystander inaction, attitudinal ambivalence theory, and general deterrence theory, laying the theoretical foundation for our study. We then introduce the research model and associated hypotheses in Section 3. In Section 4 and Section 5, we detail our research design, hypothesis testing strategy, and data analysis results. Finally, we discuss the key findings, implications, limitations, and future research directions.

2. Theoretical background

2.1. Delegitimizing doxing

Doxing is a unique form of cyberbullying that uses personal information as a weapon to harass the victim [26]. People engage in doxing for various reasons, including extortion, silencing, retribution, building reputation, and public interest [27]. Douglas [10] seminally proposed three types of doxing based on the loss caused to the target by doxing: (1) deanonymizing doxing, which reveals any identity knowledge about a person and thus removes the target's anonymity; (2) targeting doxing, which reveals information that allows the target to be physically located and thus removes the target's obscurity; and (3) delegitimizing doxing, which reveals information (e.g., moral violation and misconducts) intended to damage the target's credibility, reputation, or character. Despite different conceptualizations, three types of doxing are intertwined and usually coexist in the lifecycle of a doxing incident [10]. As discussed by Douglas [10], delegitimizing doxing often supplies evidence of a person's moral violations, which will motivate others to hunt down the person and reveal offline identifying information (e.g., work and home addresses) to target the person. That is, delegitimizing doxing supplies the motivation for targeting doxing, whereas targeting doxing serves as the means to achieve the goals of delegitimizing doxing. Moreover, delegitimizing doxing and targeting doxing will inevitably reveal the identity knowledge of the target, giving rise to deanonymizing doxing in the end.

Prior studies have examined the conceptual, moral, and legal issues of doxing [10,17,27], motivational and behavioral patterns of participants in doxing incidents [28–30], and the impact of doxing [31,32]. However, most studies do not differentiate between types of doxing, leaving conceptual ambiguity and generalizability issues. Our study focuses on delegitimizing doxing for the following reasons. First, as

compared to other types of doxing, delegitimizing doxing is more likely to go viral in that delegitimizing doxing posts usually involve moral and emotional content, which may trigger collective moral outrage and action against the target easily and thus cause disproportionate harm to the target [16,17]. Therefore, delegitimizing doxing is a more practically relevant topic that deserves more scholarly attention. Second, delegitimizing doxing is more debatable than the other two as it mostly attempts to expose moral violations to protect public values at the cost of violating the target's privacy and dignity [10,26]. As such, it is challenging for bystanders to decide whether to respond to delegitimizing doxing or not, constituting a new research opportunity for cyberbullying researchers. In this regard, our study attempts to understand bystander inaction in delegitimizing doxing.

2.2. Bystander inaction

In cyberbullying, bystanders are people who witness a cyberbullying incident [33]. Drawing upon prior studies [7,34,35], bystanders may exhibit three types of responses to cyberbullying: (1) pro-bully behavior, such as joining in the bullying and confronting the defenders of victims; (2) pro-victim behavior, such as confronting the perpetrators and reinforcers of cyberbullying, reporting cyberbullying, and comforting the victims; and (3) inaction, including ignoring cyberbullying. These responses of bystanders can shape the dynamic of cyberbullying significantly. For example, pro-bully behavior has become the invisible engine of cyberbullying as it reinforces the abuse, destroys the self-esteem of victims, signals public approval of cyberbullying, and encourages further cyberbullying against the victims [36]. In contrast, pro-victim behavior provides social support to victims and signals public disapproval of cyberbullying, reducing the harm cyberbullying causes to victims and preventing subsequent cyberbullying in the future [4].

Bystander inaction is the most common bystander response to cyberbullying because individuals are less likely to offer help to a victim in the presence of other people, as suggested by the bystander effect [21,37]. Bystander inaction has also been regarded as a form of pro-bully behavior as it may lead people to misinterpreting the cyberbullying incident as less serious than it actually is, inhibiting others' intervention attempts and encouraging perpetrators [7,37]. Many studies have sought to identify the causes of bystander inaction to inform how to motivate bystanders to stand up [5,37]. The most common determinants of bystander inaction include diffusion of responsibility, evaluation apprehension, pluralistic ignorance, lack of empathy, lack of emergency awareness, and low self-efficacy [5,33,37]. These studies have implicitly assumed bystander inaction as an instance of forced complicity and thus should be prevented [38].

However, in delegitimizing doxing, bystander inaction may help victims, rather than the perpetrators. This is because delegitimizing doxing demands a concerted effort by online and offline participants to effectively hold the target accountable and enforce social norms [15,19,30]. Other studies have also indicated that delegitimizing doxing imposes burdens on the target mainly by bringing unwanted, intense, and enduring visibility to the target [10,26]. Thus, victims of delegitimizing doxing are likely to benefit from bystander inaction, which limits their personal information and moral violations from reaching a wide audience. Moreover, as a double-edged sword that enforces social norms at the cost of violating the privacy and dignity of the target [10], delegitimizing doxing falls in a gray area where bystanders' appraisals of delegitimizing doxing can be vacillating and ambivalent. Thus, bystander inaction in delegitimizing doxing may be the result of bystanders' attitudinal ambivalence rather than other causes.

2.3. Attitudinal ambivalence

Attitudes, defined as tendencies to evaluate an attitude object positively or negatively, are regarded as one of the most robust predictors of individual behavior [39]. Most prior studies have implicitly assumed that individual attitudes toward a given object are unidimensional, either positive, negative, or neutral. However, this may ignore the richness of individuals' evaluations in reality. A growing number of researchers in various fields have challenged this unidimensional nature of attitudes and proposed the concept of attitudinal ambivalence [40,41], which is defined as the simultaneous existence of positive and negative beliefs or emotions about the same object [18]. It typically involves two components: (1) positive appraisal, which describes attraction or a pull toward an object, and (2) negative appraisal, which describes repulsion or a push away from an object [40]. Prior studies have identified multiple sources of attitudinal ambivalence, including conflicting or ambiguous norms, dynamic needs, contradictory goals, multifaceted features of attitude objects, role conflicts, and hybrid experience [40,42–44]. That is, attitudinal ambivalence usually occurs when multiple evaluation criteria are considered. In our study, by exposing alleged moral violations and personal information on SNSs, delegitimizing doxing enforces social norms at the cost of violating the privacy of the target and thus constitutes a controversial behavior. Thus, people's evaluations of delegitimizing doxing may involve multiple principles that are likely to be conflicting, resulting in attitudinal ambivalence towards delegitimizing doxing.

When people simultaneously have positive and negative evaluations of an object, they may experience cognitive dissonance and psychological discomfort and become motivated to minimize the unpleasant feelings in a variety of ways [18]. Prior studies have proposed several possible responses to attitudinal ambivalence, including compromise, holism, domination, avoidance in defense and coping mechanisms [40,45] and disengagement, denial of ambivalence, indecision, and compromise in inflexible and flexible coping responses [46,47]. In our study, we argue bystanders may respond to delegitimizing doxing by avoidance or indecision because delegitimizing doxing occurs online where bystanders are free to decide whether to intervene or not. As such, bystander inaction is a cost-effective way to reduce dissonance caused by attitudinal ambivalence in delegitimizing doxing.

2.4. General deterrence theory

General deterrence theory is rooted in rational choice theory and assumes that individuals will be deterred from a given act if the act is punished [24]. Owing to its strong theoretical foundation, general deterrence theory has been extensively applied to examine unethical or illegal behavior in various fields [48,49]. Over the years, three major features of sanctions that determine the deterrent effect have been identified in the literature [50,51]: (1) perceived severity of sanctions, which is defined as the degree to which sanctions are perceived as harsh or problematic; (2) perceived certainty of sanctions, which is defined as the degree to which sanctions are perceived as inevitable or expected; and (3) perceived celerity of sanctions, defined as the degree to which sanctions are perceived as prompt following unethical or illegal behavior. In the context of SNSs, deterrence is achieved by executing community standards and platform regulations to inform users of unacceptable and illegal behaviors. For example, regarding anti-cyberbullying, Facebook has provided clear rules about what content is allowed and prohibited and the consequences for violating community standards, such as content removal, account suspensions, or permanent bans [76]. Moreover, Facebook employs an automated system and expert reviewers to detect cyberbullying content [76]. These rules and regulations, embodying platform sanctions' severity, certainty, and celerity, deter users who might engage in cyberbullying. Following prior IS studies regarding deterrence within an organizational context (e.g., IS misuse and information security) [51,54,55], we define SNS deterrence in our study as the degree of severity, certainty, and celerity of the platform sanctions perceived by users if engaging in delegitimizing doxing.

Prior studies have identified three influencing mechanisms of how perceived deterrence of sanctions affects individuals' unethical or illegal

behavior: (1) direct effect, i.e., if people feel they will be punished severely, inevitably, and swiftly for unethical or illegal behavior, they will be deterred from this behavior to avoid risks [48,50]; (2) indirect effect, i.e., external sanctions signal the punished behavior is misaligned with moral norms in the context and thus influence individuals' behavior indirectly by changing their moral evaluations of the behavior [52,53]; and (3) moderating effect, i.e., external sanctions serve as environmental conditions under which personal factors influence individuals' behavior [54,55]. Given that more and more studies adopt interactionism to portray unethical or illegal behavior as the product of the interplay between personal and environmental factors [25], we consider perceived SNS deterrence as a moderator to offer a more integrative view of how attitudinal ambivalence and perceived SNS deterrence collectively affect bystander inaction in delegitimizing doxing.

3. Research model and hypothesis development

Fig. 1 depicts our research model explaining bystander inaction in delegitimizing doxing. In line with the unique nature of delegitimizing doxing, we apply attitudinal ambivalence theory and posit that bystander inaction in delegitimizing doxing is the result of bystanders' positive and negative appraisals of delegitimizing doxing at the same time. To better articulate how attitudinal ambivalence affects bystander inaction, we follow prior studies of congruence and hypothesize three types of congruence effects [56,57]: (1) the effect of congruence types, describing how bystander inaction varies as the level of positive and negative appraisal increases in the same direction; (2) the effect of the degree of incongruence, describing how bystander inaction varies as the level of positive and negative appraisal increases in the opposite direction; and (3) the effect of the incongruence types, describing how the directions of incongruence between positive and negative appraisal relate to bystander inaction. Fig. 2 illustrates the three types of congruence effects.

Furthermore, drawing upon the interactionism that highlights the importance of person-environment interaction in explaining unethical or illegal behavior [21], we apply general deterrence theory to understand the environmental conditions of bystander inaction and argue attitudinal ambivalence and perceived SNS deterrence will collectively influence bystander inaction in delegitimizing doxing. Finally, to ensure the robustness of our model, we control several demographic variables (e.g., age, gender, education, social media usage experience), common determinants of bystander inaction (e.g., perceived responsibility, perceived empathy), and the fixed effect of the types of delegitimizing doxing in the model.

3.1. Attitudinal ambivalence and bystander inaction

We first hypothesize how bystanders' attitudinal ambivalence in different levels of positive and negative appraisals of delegitimizing doxing affects their inaction (i.e., how bystander inaction varies across Quadrant I and Quadrant III in Fig. 2). In the case where bystanders' appraisals of delegitimizing doxing are neither positive nor negative (Quadrant III), delegitimizing doxing may fail to attract bystanders' attention and as a result, bystanders are less likely to be emotionally triggered by delegitimizing doxing [40]. That is, bystanders will become indifferent to delegitimizing doxing and will not be motivated to respond to delegitimizing doxing. The low levels of positive and negative appraisals of delegitimizing doxing may also lead bystanders to believe delegitimizing doxing is not urgent, thereby increasing the likelihood of bystander inaction in delegitimizing doxing [5]. In the case where bystanders' positive and negative appraisals of delegitimizing doxing are both high (Quadrant I), the mixed and conflicting appraisals may lead bystanders to suffer cognitive dissonance and then behaviorally disengage from delegitimizing doxing to avoid discomfort [47]. When bystanders have both positive and negative appraisals of delegitimizing doxing, they are uncertain about their attitudes towards delegitimizing doxing and become vacillating, thus delaying their decisions in responding to delegitimizing doxing [47,58]. Overall, bystanders tend to stand by in delegitimizing doxing when their positive and negative appraisals of delegitimizing doxing are high or low, indicating no significant difference in bystander inaction in delegitimizing doxing. Thus, we hypothesize:

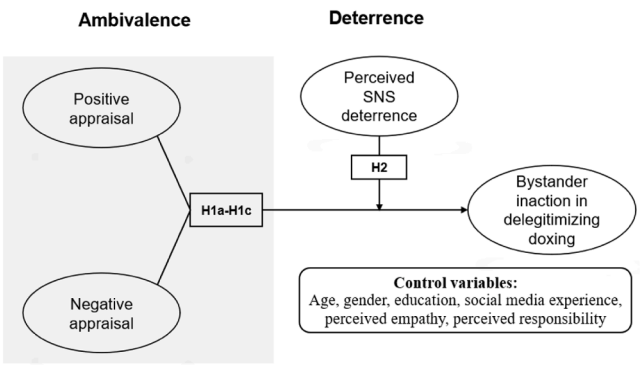


Fig. 1. Research Model.

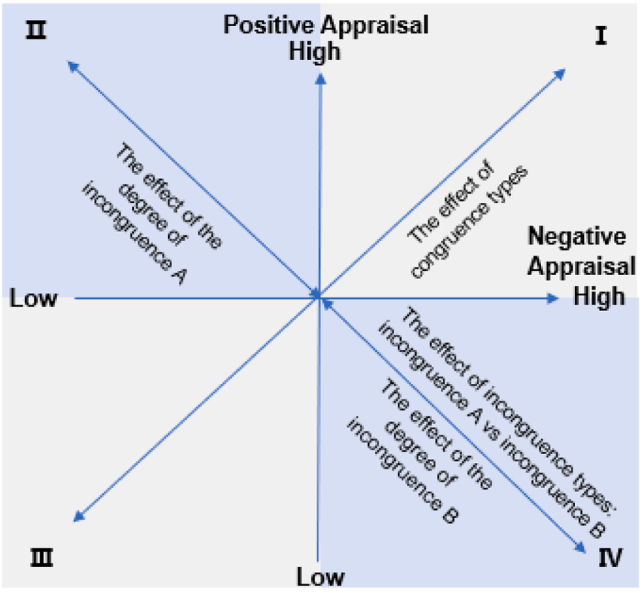


Fig. 2. Illustration of congruence effects.

H1a: Attitudinal ambivalence in different levels of positive and negative appraisals of delegitimizing doxing does not affect bystander inaction in delegitimizing doxing, such that as positive and negative appraisals increase in the same direction, the likelihood of bystander inaction in delegitimizing doxing remains unchanged.

We further hypothesize how the incongruence between bystanders' positive and negative appraisals of delegitimizing doxing relates to bystander inaction in delegitimizing doxing. As illustrated in Fig. 2, there are two cases of incongruence between bystanders' positive and negative appraisals of delegitimizing doxing: (1) positive attitudinal univalence (Quadrant II), where bystanders' positive appraisal of delegitimizing doxing is high whereas their negative appraisal of delegitimizing doxing is low; and (2) negative attitudinal univalence (Quadrant IV), where bystanders' negative appraisal of delegitimizing doxing is high, whereas their positive appraisal of delegitimizing doxing is low. As elaborated in H1a, bystander inaction is the result of the same levels

(either high or low) of positive and negative appraisals of delegitimizing doxing. Accordingly, once bystanders' positive and negative appraisals are incongruent, they are less likely to stand by in delegitimizing doxing. In the case of positive attitudinal univalence (Quadrant II), bystanders' positive appraisals of delegitimizing doxing outweigh their negative appraisals and thus, their attitude towards delegitimizing doxing will be unambiguously positive. As such, bystanders will be motivated to join in delegitimizing doxing to take justice into their own hands and contribute to delegitimizing doxing, reducing the likelihood of bystander inaction in delegitimizing doxing. In the case of negative attitudinal univalence (Quadrant IV), bystanders' negative appraisals of delegitimizing doxing outweigh their positive appraisals. In this case, bystanders may treat delegitimizing doxing as an emergent issue and thus be motivated to intervene in delegitimizing doxing, rather than standing by [5]. Overall, as the incongruence between bystanders' positive and negative appraisals increases, despite different responses, bystanders are less likely to do nothing in delegitimizing doxing. Therefore, we hypothesize:

H1b: Attitudinal univalence negatively affects bystander inaction in delegitimizing doxing, such that as the incongruence between positive and negative appraisals of delegitimizing doxing increases, the likelihood of bystander inaction in delegitimizing doxing decreases.

Following prior studies of congruence [56,59], we further hypothesize the negative effects of two forms of incongruence (positive and negative attitudinal univalence) on bystander inaction are asymmetrical. As elaborated in H1b, bystanders are likely to be triggered to join in delegitimizing doxing to enforce social norms and signal moral quality if their positive appraisals outweigh negative appraisals of delegitimizing doxing, indicating a lower likelihood of bystander inaction in delegitimizing doxing. In contrast, when bystanders' negative appraisals outweigh positive appraisals of delegitimizing doxing, they may help the victims because of responsibility and empathy, rather than stand by. However, according to the bystander intervention model, bystander intervention (e.g., reporting, comforting the victims, confronting the perpetrators) involves more information processing and thus requires stronger motivation of bystanders [5,21]. In particular, victims of delegitimizing doxing are often portrayed as norm transgressors [10], in which case bystanders are instinctively reluctant to offer help to avoid any reputational risk. This line of reasoning suggests that although both positive and negative attitudinal univalence may reduce bystander inaction in delegitimizing doxing, bystanders are less likely to stand by in the case of positive attitudinal univalence than in the case of negative attitudinal univalence. Therefore, we hypothesize:

H1c: Positive attitudinal univalence has a stronger negative impact on bystander inaction in delegitimizing doxing than negative attitudinal univalence, such that bystanders are less likely to stand by in the case of positive attitudinal univalence than in the case of negative attitudinal univalence.

3.2. Moderating role of perceived SNS deterrence

According to interactionism, individuals' behavior, especially unethical or illegal behavior, is the interactive product of the individual and the situation [25]. We thus posit that perceived SNS deterrence (environmental factors) will shape the effects of attitudinal ambivalence (personal factors) on bystander inaction in delegitimizing doxing. Given that bystander inaction remains unchanged as two equal levels of positive and negative appraisals increase (H1a), we only hypothesize how perceived SNS deterrence moderates the effect of attitudinal univalence on bystander inaction (H1b). As mentioned, bystanders may join in delegitimizing doxing to take justice into their own hands and contribute to delegitimizing doxing because of positive attitudinal univalence. When perceived SNS deterrence is high, suggesting that delegitimizing doxing will be effectively regulated and punished on the platforms, the risks and costs to engage in delegitimizing doxing will be high. In other words, on SNSs with high deterrence, bystanders who violate community standards (e.g., joining in delegitimizing doxing) are more likely to be severely punished by platforms, such as content removal, account suspensions, or permanent bans. Therefore, the bystanders' intention to join in delegitimizing doxing because of positive attitudinal univalence is likely to be weakened, i.e., the likelihood of bystander inaction is less likely to be reduced in this case. This is consistent with a large body of studies showing that external sanctions will cause people to rationally reduce engagement in unethical or illegal behavior they were originally motivated to commit [25,36]. Moreover, the high level of perceived SNS deterrence suggests that the enforcement of external sanctions against delegitimizing doxing is effective, which may cause bystanders to underestimate the necessity of their intervention in delegitimizing doxing in the case of negative attitudinal univalence. This is because external sanctions can be considered as alternatives to bystander intervention in the prevention of delegitimizing doxing. Prior studies have shown that the attractiveness or effectiveness of alternative solutions would inhibit people from engaging in the focal solution [60,61]. Overall, in the two cases of attitudinal univalence, bystander inaction is less likely to be reduced when perceived SNS deterrence is high than when perceived SNS deterrence is low. Therefore, we hypothesize:

H2: Perceived SNS deterrence weakens the negative effect of attitudinal univalence on bystander inaction in delegitimizing doxing.

4. Research method

4.1. Data collection and samples

To validate the research model and associated hypotheses, we used an online survey to collect data by collaborating with a third-party crowdsourcing platform. This platform provides us with access to participants across demographics and backgrounds at minimal recruitment costs. We also followed the guidelines proposed by several methodological studies to ensure the proper use of online crowdsourcing platforms in data collection [62,63]. We first provided a working definition of delegitimizing doxing at the outset of the survey to ensure participants had a consistent understanding of delegitimizing doxing. Second, we asked participants to describe details of a delegitimizing doxing incident they had seen on social media platforms, including the doxer, the doxee, the relationship between them, the moral violation and associated evidence, and the personal information disclosed. Participants reported a wide range of delegitimizing doxing incidents, including exposing private sexual relationships, intimate partner violence, illegal business activities, corruption of officials, animal abuse, uncivilized public behaviors, teacher-student conflicts, and hypocrisy of celebrities. Third, we also asked participants to provide the social media platform where they saw the delegitimizing doxing post. The platforms provided included WeChat, Weibo, TikTok, Kuaishou, Douban, and Baidu Post, among others. Attitudinal ambivalence and SNS deterrence constructs were measured based on participants' responses to the reported delegitimizing doxing incident and the social media platform. As such, we could observe a wide range of variances in participants' perceptions of attitudinal ambivalence and SNS deterrence constructs. To ensure the response quality, we strategically designed several attention-check questions and filtered out participants with lowered reputation ratings, suspiciously robotic responses, and multiple repeated responses by using the platform's features. Finally, we obtained a total of 542 valid responses. Table 1 details the profiles of samples.

We further conducted an exploratory open-ended survey to explore how an individual forms an ambivalent attitude toward delegitimizing doxing. Drawing upon the teleological (i.e., results-based evaluations) and deontological (i.e., rule-based evaluations) perspectives [79], we

Table 1
Profiles of samples.

Characteristic	N	%	Characteristic	N	%
Gender			**Age**		
Male	238	43.91	18–24	183	33.76
Female	304	56.08	25–30	164	30.26
Education			31–40	155	28.60
High school (equivalent) or below	13	2.40	41–54	32	5.90
Bachelor's (equivalent)	448	82.66	≥ 55	8	1.48
Master's or above	81	14.94	**Experience using the platform**		
Time spent on the platform every day			≤6 months	2	0.37
≤ 30 min	23	4.24	6–12 months	5	0.92
30–60 min	131	24.17	1–2 years	22	4.06
1–2 h	226	41.70	3–5 years	215	39.67
3–5 h	130	23.99	6–10 years	214	39.48
> 5 h	32	5.90	≥10 years	84	15.50

asked participants to comment on the positive and negative consequences of delegitimizing doxing and the social norms that support and oppose delegitimizing doxing, respectively. The qualitative responses showed that every participant formed an ambivalent attitude toward delegitimizing doxing as they reported a wide range of positive and negative consequences of delegitimizing doxing and the social norms that support and oppose delegitimizing doxing. Specifically, positive consequences of delegitimizing doxing include exposing misconduct, upholding social justice, informing the public, and fostering network consensus supervision, whereas negative consequences of delegitimizing doxing include disproportionate retaliation against the doxee, harm to their family members, unintended harm to innocent individuals, and hindrances to law enforcement efforts. The social norms supporting delegitimizing doxing include duties to protect public values, whereas social norms opposing delegitimizing doxing include violations of privacy and dignity. The results reveal a prevalent sense of ambivalence among individuals towards delegitimizing doxing practices, offering preliminary support for our research model.

4.2. Operationalization of constructs

Attitudinal Ambivalence. Two approaches have been used to operationalize attitudinal ambivalence [18]: (1) direct measures, which use items expressing the subjective feeling of ambivalence; and (2) indirect measures, which assess separate positive and negative evaluations. Due to the methodological issues of direct measures (e.g., conceptual ambiguity, discarded information, recall bias, debatable assumption) [18,57,64], we used indirect measures to operationalize attitudinal ambivalence. This is also directly related to the definition of ambivalence as the coexistence of positive and negative evaluations of the same object [18]. We first asked participants to evaluate the positive qualities of delegitimizing doxing while ignoring its negative qualities. Then we asked them to evaluate the negative qualities of delegitimizing doxing while ignoring its positive ones. Items measuring positive and negative appraisals were adapted from [65].

Perceived SNS Deterrence. We operationalized perceived SNS deterrence by using perceived severity, certainty, and celerity of sanctions. Items for measuring perceived severity, certainty, and celerity of sanctions were adapted from [50,51]. Drawing upon studies of higher-order constructs [66], we considered perceived SNS deterrence as a second-order formative construct because perceived severity, certainty, and celerity of sanctions would cumulatively contribute to the overall deterrent effect of sanctions.

Bystander Inaction. We used items adapted from [34,35] to measure bystander inaction.

We summarized all items and prompts in Table 2.

Table 2
Measurement items.

Constructs	Items
Positive appraisal (PA)	Think about your evaluation of the post exposing personal information and moral violation of the target and answer the following questions: **PA1:** Considering only the favorable aspects of the post and ignoring its unfavorable aspects, how favorable is your evaluation of the post? ("1→7": "Not at all favorable → Extremely favorable") **PA2:** Considering only the positive aspects of the post and ignoring its negative aspects, how positive is your evaluation of the post? ("1→7": "Not at all positive → Extremely positive") **PA3:** Considering only the beneficial aspects of the post and ignoring its harmful aspects, how beneficial is your evaluation of the post? ("1→7": "Not at all beneficial → Extremely beneficial")
Negative appraisal (NA)	Think about your evaluation of the post exposing personal information and moral violation of the target and answer the following questions: **NA1:** Considering only the unfavorable aspects of the post and ignoring its favorable aspects, how unfavorable is your evaluation of the post? ("1→7": "Not at all unfavorable → Extremely unfavorable") **NA2:** Considering only the negative aspects of the post and ignoring its positive aspects, how negative is your evaluation of the post? ("1→7": "Not at all negative → Extremely negative") **NA3:** Considering only the harmful aspects of the post and ignoring its beneficial aspects, how harmful is your evaluation of the post? ("1→7": "Not at all harmful → Extremely harmful")
Perceived severity of sanctions (PSS)	Think about your evaluation of the social media platform you provide and indicate to what extent you agree with the following statements ("1→7": "Strongly Disagree → Extremely agree"). **PSS1:** The platform's sanctions on users for delegitimizing doxing posts that expose others' personal information and moral violations are severe. **PSS2:** The platform will harshly punish users for delegitimizing doxing posts that expose others' personal information and moral violations. **PSS3:** The platform will take strict actions against users for delegitimizing doxing posts that expose others' personal information and moral violations.
Perceived certainty of sanctions (PCT)	Think about your evaluation of the social media platform you provide and indicate to what extent you agree with the following statements ("1→7": "Strongly Disagree → Extremely agree"). **PCT1:** The likelihood that the platform will punish users for delegitimizing doxing posts that expose others' personal information and moral violations is high. **PCT2:** The platform is very likely to punish users for delegitimizing doxing posts that expose others' personal information and moral violations. **PCT3:** Users are probably punished by the platform for delegitimizing doxing posts that expose others' personal information and moral violations.
Perceived celerity of sanctions (PCL)	Think about your evaluation of the social media platform you provide and indicate to what extent you agree with the following statements ("1→7": "Strongly Disagree → Extremely agree"). **PCL1:** The platform's sanctions on users for delegitimizing doxing posts that expose others' personal information and moral violations are timely. **PCL2:** The platform will swiftly punish users for delegitimizing doxing posts that expose others' personal information and moral violations. **PCL3:** Users are quickly punished by the platform for delegitimizing doxing posts that expose others' personal information and moral violations.

(*continued on next page*)

Table 2 (*continued*)

Constructs	Items
Bystander inaction in delegitimizing doxing (BI)	If I witness a similar delegitimizing doxing post that exposes others' personal information and moral violations again, the likelihood that I engage in the following behaviors is ("1→7": "Absolutely impossible → Absolutely possible"). **BI1:** I choose to stay out of it. **BI2:** I ignore it. **BI3:** I do not do anything but wait and see what happens. **BI4:** I don't support any party involved in the incident.

5. Data analysis and hypothesis testing results

5.1. Construct reliability and validity

Before evaluating construct reliability and validity, we first evaluate the risk of common method bias (CMB) using several methods proposed by [67]. First, we performed Harman's single-factor test, and the results indicated that the extracted primary component accounted for less than half of the variance in the data (36.964 %). Second, drawing upon [77], we employed a one-factor latent model that associated all items with a single factor. The results of one-factor confirmatory factor analysis (CFA) demonstrated extremely unsatisfactory model fit indices: $\chi 2$ [152] = 4169.202, $p < 0.001$, RMSEA=0.221, CFI=0.530, TLI=0.471, SRMR=0.192, and CD=0.962, indicating negligible CMB issues. Third, we followed the guideline of [78] and ran a CFA model by adding a common methods factor that includes all items. By comparing with our measurement model, the results demonstrated that CMB was not a serious concern in our study: ΔX2/df =0.111, ΔRMSEA=0.003, ΔCFI=−0.003, ΔTLI=−0.003, ΔSRMR=0.002, ΔCD=0.

We then performed the CFA in Stata 14.0 using the covariance-based structural equation modeling to assess the reliability and validity of the first-order constructs. The CFA results demonstrated the model fit indices were acceptable: $\chi 2$ [137] = 236.657, $p < 0.001$, RMSEA=0.037, CFI=0.988, TLI=0.985, SRMR=0.034, and CD=1.000. Moreover, as shown in Table 3, item loadings for each construct were above the threshold of 0.7, and the average variance extracted (AVE) values were all greater than 0.5, indicating good convergent validity for each construct. Cronbach's alpha values ranged between 0.872 and 0.947, whereas composite reliability values ranged between 0.873 and 0.947, exceeding the threshold of 0.7. Therefore, all constructs were reliable. As shown in Table 4, the square root of the AVE value for each construct was higher than the correlation between that construct and all other constructs, indicating good discriminant validity for each construct. The variance inflation factors range from 1.11 to 3.79, far below the threshold of 10, thus indicating multicollinearity is not a threat to our study.

Table 3
Construct reliability and validity.

Construct	Mean	SD	CA	CR	AVE	Items	Loading
Positive appraisal (PA)	5.429	1.152	0.880	0.881	0.713	PA1 PA2 PA3	0.863 0.841 0.828
Negative appraisal (NA)	4.261	1.453	0.872	0.873	0.697	NA1 NA2 NA3	0.754 0.881 0.865
Perceived severity of sanctions (PSS)	3.690	1.688	0.947	0.947	0.857	PSS1 PSS3 PSS3	0.910 0.943 0.924
Perceived certainty of sanctions (PCT)	4.149	1.633	0.936	0.937	0.831	PCT1 PCT2 PCT3	0.890 0.916 0.929
Perceived celerity of sanctions (PCL)	3.798	1.561	0.923	0.925	0.804	PCL1 PCL2 PCL3	0.851 0.926 0.911
Bystander inaction (BI)	4.024	1.539	0.887	0.890	0.670	BI1 BI2 BI3 BI4	0.881 0.781 0.883 0.717

Note. SD: standard deviation; CA: Cronbach's alpha; CR: composite reliability; AVE: average variance extracted.

5.2. Hypothesis testing strategy

Prior studies of attitudinal ambivalence relying on indirect measures used diverse formulae that combined positive and negative evaluations arithmetically to compute a single ambivalence index [18,65]. However, as discussed by [18], most existing formulae failed to fulfill all the conditions proposed by [68] for creating an ideal ambivalence index. Worse, prior studies have also extensively criticized the methodological issues of a single-index measure [64,69]. To avoid conceptual and methodological issues, we followed prior studies of congruence and leveraged PRRSA to test the effect of attitudinal ambivalence (H1a-H1c) [56,59,69]. Instead of collapsing measures of the two predictor variables into a single index, PRRSA estimated the main, interaction, and nonlinear effects of the two variables, as shown in Eq. (1). The coefficients ($b_1 \sim b_5$) were then used to plot a three-dimensional response space, and the features of the response surface were tested to uncover how combinations of two predictors relate to the outcome variable. By doing so, PRRSA overcomes the limitations of single-index measures and has been increasingly used in the literature [57,59,70].

$$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + e \quad (1)$$

In Eq. (1), Z is bystander inaction in delegitimizing doxing, X is positive appraisal, and Y is negative appraisal in our study. According to prior studies of congruence [56,71], the key response surface features include: (1) the line of congruence (LOC) ($X = Y$), and the slope (a_1) and curvature (a_2) of the surface along the LOC can be used to understand how the outcome variable varies when two predictor variables are equal and increase in the same direction; (2) the line of incongruence (LOIC) ($X = -Y$), and the slope (a_3) and curvature (a_4) of the surface along the LOC can be used to understand how the outcome variable varies as the degree of the discrepancy between two predictor variables increases; and (3) the first principal axis (FPA), and the intercept (p_{10}) and the slope (p_{11}) of the FPA can be used to understand whether the outcome variable varies symmetrically in the two cases of incongruence [59,70]. The slopes and curvatures of the surface along the LOC and LOIC ($a_1 \sim a_4$) and their significance levels ($t_1 \sim t_4$) can be obtained by the following equations:

$$a_1 = b_1 + b_2 \quad (2)$$

$$t_1 = a_1 \Big/ \sqrt{(SE_{b1}^2 + SE_{b2}^2) + 2COV_{b1b2}} \quad (3)$$

$$a_2 = b_3 + b_4 + b_5 \quad (4)$$

$$t_2 = a_2 \Big/ \sqrt{(SE_{b3}^2 + SE_{b4}^2 + SE_{b5}^2) + 2COV_{b3b4} + 2COV_{b3b5} + 2COV_{b4b5}} \quad (5)$$

$$a_3 = b_1 - b_2 \quad (6)$$

$$t_3 = a_3 \Big/ \sqrt{(SE_{b1}^2 + SE_{b2}^2) - 2COV_{b1b2}} \quad (7)$$

$$a_4 = b_3 - b_4 + b_5 \quad (8)$$

Table 4
Correlation matrix.

Construct	VIF	PA	NA	PSS	PCT	PCL	BI
PA	1.25	**0.844**					
NA	1.28	−0.428***	**0.835**				
PS	3.79	−0.197***	0.190***	**0.926**			
PCT	3.32	−0.217***	0.262***	0.823***	**0.912**		
PCL	2.18	−0.119*	0.135*	0.725***	0.664***	**0.897**	
BI	1.11	−0.056	0.208***	−0.131*	−0.019	−0.160**	**0.819**

Note. *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$. BI: bystander inaction; NA: negative appraisal; PA: positive appraisal; PCL: perceived celerity of sanctions; PCT: perceived certainty of sanctions; PSS: perceived severity of sanctions; VIF: variance inflation factor.

$$t_4 = a_4 \Big/ \sqrt{(SE_{b3}^2 + SE_{b4}^2 + SE_{b5}^2) - 2COV_{b3b4} + 2COV_{b3b5} - 2COV_{b4b5}} \tag{9}$$

where $b_1 \sim b_5$ are unstandardized coefficients estimated by Eq. (1), $SE_{b1} \sim SE_{b5}$ are standard errors of $b_1 \sim b_5$, and COV is the covariance between the corresponding coefficients. The intercept (p_{10}) and the slope (p_{11}) of the FPA can be given by the following equations:

$$p_{10} = Y_0 - p_{11}X_0 \tag{10}$$

$$p_{11} = \frac{b_5 - b_3 + \sqrt{(b_3 - b_5)^2 + b_4^2}}{b_4} \tag{11}$$

$(X_0,\ Y_0)$ is the stationary point of the response surface [69]. The values of X_0 and Y_0 can be given by the following equations:

$$X_0 = \frac{b_2 b_4 - 2b_1 b_5}{4b_3 b_5 - b_4^2} \tag{12}$$

$$Y_0 = \frac{b_1 b_4 - 2b_2 b_3}{4b_3 b_5 - b_4^2} \tag{13}$$

To test whether the incongruence effects are symmetrical, researchers first need to ensure that the effect of the degree of the discrepancy between the two predictor variables on the outcome variable is significant ($a_4 \neq 0$). Then, researchers need to check whether the FPA is rotated or laterally shifted off the LOC, i.e., whether $p_{10} = 0$ and $p_{11} = 1$ [59,70], in which case a bootstrapping approach has been widely used [72].

Furthermore, to test the moderating effect of perceived SNS deterrence (H2), we followed prior studies and performed moderated polynomial regression analysis using Eq. (14) [73]. In our study, W is the moderator, i.e., perceived SNS deterrence. The moderated congruence hypotheses were tested by comparing the response surface features under the values of W at one standard deviation above versus below its mean [73].

$$Z = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 + b_6W + b_7XW + b_8YW + b_9X^2W + b_{10}XYW + b_{11}Y^2W + e \tag{14}$$

The key features of the response surface under moderated polynomial regression analysis (a_x, a_{x^2}, a_y, and a_{y^2}) can be given by the following equations. We also used the bootstrapping approach to test their significance levels [73].

$$a_x = b_1 + b_2 + (b_7 + b_8)W \tag{15}$$

$$a_{x^2} = b_3 + b_4 + b_5 + (b_9 + b_{10} + b_{11})W \tag{16}$$

$$a_y = b_1 - b_2 + (b_7 - b_8)W \tag{17}$$

$$a_{y^2} = b_3 - b_4 + b_5 + (b_9 - b_{10} + b_{11})W \tag{18}$$

Drawing upon several well-established studies of PRRSA [59,69,71,73], we summarized the tests of response surface features required for each hypothesis in Table 5.

5.3. Hypothesis testing results

Before performing PRRSA, we first scale-centered all items of independent variables to control the multicollinearity and increase the interpretability of the response surface [57,69]. To perform PRRSA, all first-order constructs were measured by averaging their items, whereas perceived SNS deterrence was measured by averaging factor scores of its subconstructs (perceived severity, certainty, and celerity of sanctions), as suggested by prior studies of second-order formative constructs [74]. We control the fixed effects of types of delegitimizing doxing in our model. Table 6 shows the PRRSA results. Fig. 3 depicts the response surface of polynomial regression analysis results. As shown in Table 6,

Table 5
Hypothesis testing strategies.

Hypothesis	Description and Test
H1a: Attitudinal ambivalence in different levels of positive and negative appraisals of delegitimizing doxing does not affect bystander inaction in delegitimizing doxing, such that as positive and negative appraisals increase in the same direction, the likelihood of bystander inaction in delegitimizing doxing remains unchanged.	Bystander inaction remains unchanged along the LOC. That is, the response surface along the LOC is flat. **Test 1:** $a_1 = 0$; and **Test 2:** $a_2 = 0$.
H1b: Attitudinal univalence negatively affects bystander inaction in delegitimizing doxing, such that as the incongruence between positive and negative appraisals of delegitimizing doxing increases, the likelihood of bystander inaction in delegitimizing doxing decreases.	Bystander inaction decreases as either positive attitudinal univalence or negative attitudinal univalence increases, indicating an inverted-U shape of the response surface along the LOIC. **Test 3:** $a_4 < 0$.
H1c: Positive attitudinal univalence has a stronger negative impact on bystander inaction in delegitimizing doxing than negative attitudinal univalence, such that bystanders are less likely to stand by in the case of positive attitudinal univalence than in the case of negative attitudinal univalence.	Bystander inaction decreases faster in the case of positive attitudinal univalence than in the case of negative attitudinal univalence, indicating the surface is laterally shifted or rotated counterclockwise toward the region where $Y > X$ (i.e., negative attitudinal univalence in this study). **Test 4:** The intercept of the FPA $p_{10} > 0$ and lateral shift quantity is negative; or **Test 5:** The intercept of the FPA $p_{10} = 0$ and the slope of the FPA $p_{11} > 1$.
H2: Perceived SNS deterrence weakens the negative effect of attitudinal univalence on bystander inaction in delegitimizing doxing.	When perceived SNS deterrence is low, bystander inaction in delegitimizing doxing decreases faster in two cases of attitudinal univalence than when perceived SNS deterrence is high. **Test 6:** $a_{y^2} < 0$; and **Test 7:** The absolute value of the a_{y^2} is greater when perceived SNS deterrence is low than when perceived SNS deterrence is high.

Table 6
Results of polynomial regression analysis.

Constructs	Model 1	Model 2	Model 3
Intercept	2.794**	4.183**	5.051***
Gender	0.315*	0.191	0.171
Age	−0.240***	−0.165**	−0.144*
Education	0.376**	0.326*	0.304*
Social media experience	−0.047	−0.024	−0.038
Perceived empathy	0.102	0.068	0.077
Perceived responsibility	−0.201**	−0.192**	−0.148*
Positive appraisal (PA) (b_1)	−0.014	0.012	−0.114
Negative appraisal (NA) (b_2)	0.252***	0.076	−0.015
PA*PA (b_3)		−0.161***	−0.112*
PA*NA (b_4)		0.088	0.133*
NA*NA (b_5)		−0.029	−0.026
Perceived SNS deterrence (PD) (b_6)			−0.162
PA*PD (b_7)			0.132
NA*PD (b_8)			0.068
PA*PA*PD (b_9)			−0.033
PA*NA*PD (b_{10})			−0.050
NA*NA*PD (b_{11})			−0.017
Line of congruence			
Slope: a_1		0.089	
Curvature: a_2		−0.102	
Line of incongruence			
Slope: a_3		−0.064	
Curvature: a_4		−0.279***	
First principal axis			
Intercept: p_{10}		0.103	
Slope: p_{11}		3.285	
FE (types of delegitimizing doxing)	Yes	Yes	Yes
R^2	0.149	0.214	0.225
ΔR^2		0.065	0.011

Note. *$p < 0.05$, ***$p < 0.01$, ***$p < 0.001$. Fe: fixed effects. Dependent variable is bystander inaction in delegitimizing doxing.

the slope and curvature of the response surface along the LOC were not significant ($a_1 = 0.089$, $p > 0.05$; $a_2 = -0.102$, $p > 0.05$), supporting Test 1 and Test 2 in Table 5. That is, bystander inaction in delegitimizing doxing remained unchanged along the LOC. Therefore, **H1a** was supported. The curvature of the response surface along the LOIC was significant and negative ($a_4 = -0.279$, $p < 0.001$), supporting Test 3 in Table 5. That is, bystander inaction in delegitimizing doxing decreases as either positive or negative attitudinal univalence increases. Therefore, **H1b** was supported.

Moreover, the intercept of the FPA p_{10} was 0.103, and its bias-corrected 95 % confidence interval (BC 95 % CI) after bootstrapping 5000 resamples included 0. That is, the intercept of the FPA was not significantly different from 0, indicating no lateral shift of the FPA from the LOC [71,72]. Thus, Test 4 was not supported. The slope of the FPA p_{11} was 3.285, and its BC 95 % CI after bootstrapping 5000 resamples excluded 1. As such, p_{11} was significantly greater than 1, indicating a counterclockwise rotation off the LOC [71,72]. Thus, Test 5 was supported. Overall, the results ($p_{10} = 0$, $p_{11} > 1$) showed that the surface was rotated counterclockwise towards the region where $Y > X$ (i.e., negative attitudinal univalence in our study), and as a result, bystander inaction in delegitimizing doxing decreased faster in the case of positive attitudinal univalence than in the case of negative attitudinal univalence. Therefore, **H1c** was supported.

Table 7 shows the results of moderated polynomial regression analysis. As shown in Table 7, Test 6 and Test 7 in Table 5 were supported: regardless of the level of perceived SNS deterrence, the curvatures of the response surface along the LOIC were negative and significant, with BC 95 % CIs excluding 0. Moreover, $a_{y^2 High} > a_{y^2 Low}$. Thus, H2 was supported.

6. Discussion and implications

6.1. Discussion of key findings

This study aims to understand bystander inaction in delegitimizing doxing. To explore this, we combined attitudinal ambivalence and general deterrence theory to develop an integrated model to explain bystander inaction in delegitimizing doxing. Using PRRSA to analyze the effects of attitudinal ambivalence, we uncover three distinctive congruence effects describing how different combinations of positive

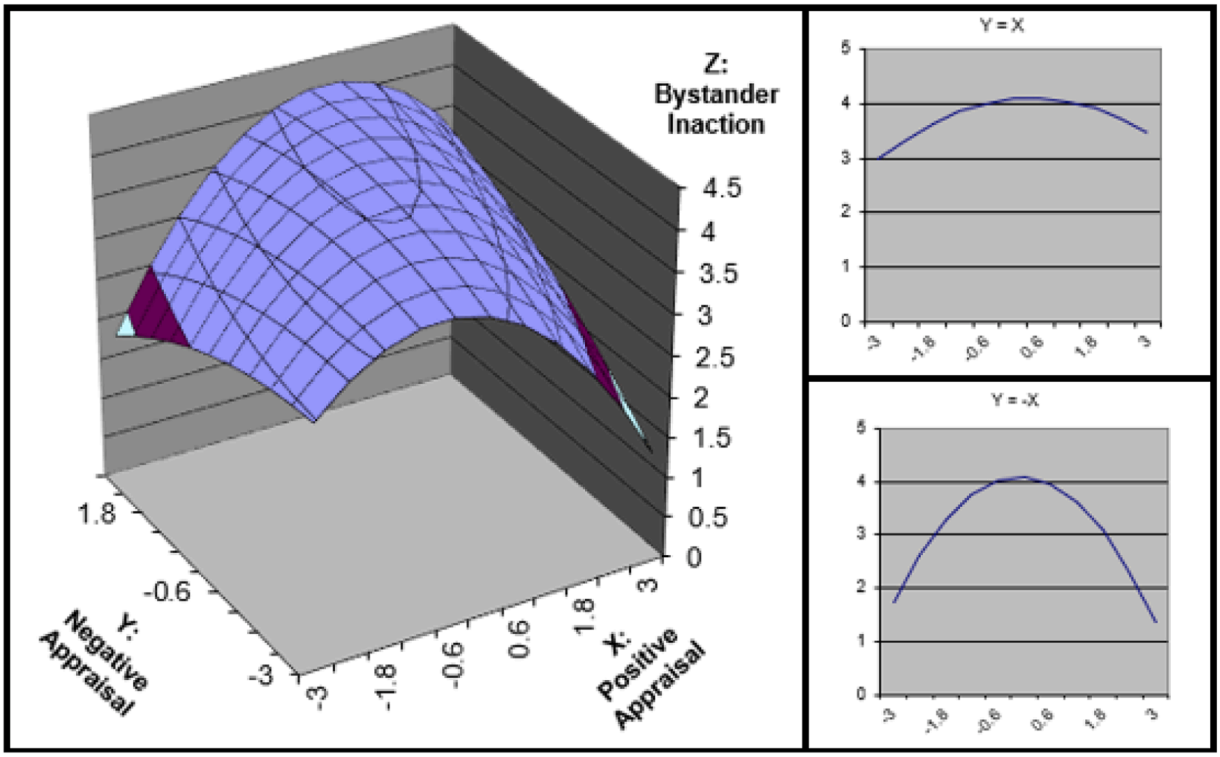


Fig. 3. Response surface for attitudinal ambivalence and bystander inaction.

Table 7
Results of moderated polynomial regression analysis.

Level of perceived SNS deterrence	Line of congruence		Line of incongruence	
	Slope a_x	Curvature a_{x^2}	Slope a_y	Curvature a_{y^2}
High	0.142	−0.139	−0.014	−0.269
[BC 95 % CI]	[−0.400, 0.676]	[−0.349, 0.107]	[−0.173, 0.140]	[−0.373, −0.165]
Low	−0.451	0.157	−0.200	−0.271
[BC 95 % CI]	[−1.096, 0.199]	[−0.139, 0.453]	[−0.445, 0.056]	[−0.396, −0.148]

Note. BC 95 % CI: bias-corrected 95 % confidence interval, with 5000 bootstrapping iterations.

and negative appraisals determine bystander inaction in delegitimizing doxing. First, we find that bystander inaction in delegitimizing doxing remains unchanged when bystanders' positive and negative appraisals of delegitimizing doxing are equal and increase in the same direction (i.e., the effect of congruence types). Different from the well-established ambivalence response framework [40], we demonstrate that bystanders may respond to ambivalence by using avoidance rather than compromise and holism when both positive and negative appraisals are both medium and high, thus challenging prior studies [40]. One of the possible reasons for this is that delegitimizing doxing is an online phenomenon that is easy to escape for users and forms their free willingness to respond, which may thus facilitate them to adopt an avoidance strategy to reduce unpleasant feelings caused by ambivalence.

Second, consistent with the ambivalence response framework [40], our study demonstrates that both positive and negative attitudinal univalence limits bystander inaction in delegitimizing doxing (i.e., the effect of the degree of incongruence). Specifically, bystanders are less likely to stand by in delegitimizing doxing, as the degree of incongruence between positive and negative appraisals increases. In such cases, bystanders tend to adopt the domination strategy, a frequent response adopted by individuals when positive or negative evaluations are over the other [40,43]. We thus infer that positive attitudinal univalence motivates bystanders to join in delegitimizing doxing, whereas negative attitudinal univalence motivates bystanders to help the victims, hence reducing bystander inaction in delegitimizing doxing.

Third, we also find that positive attitudinal univalence has a stronger negative effect on bystander inaction than negative attitudinal univalence (i.e., the effect of the incongruence types). It suggests bystanders are more likely to reduce inaction in the case where positive appraisals outweigh negative appraisals than in the case where negative appraisals are stronger than positive appraisals. Thus, we provide a possible explanation why delegitimizing doxing is often a hot topic, in which bystanders tend to impulsively join in because of their positive attitudinal univalence.

Last, our study provides empirical evidence of the importance of considering SNS deterrence against delegitimizing doxing. We demonstrate that SNS deterrence weakens the negative impact of positive and negative attitudinal univalence on bystander inaction. It suggests that when perceived SNS deterrence is high, bystanders' intention to join in delegitimizing doxing because of positive attitudinal univalence will be lowered to avoid risks and costs of being punished, whereas bystanders' intention to help the victims because of negative attitudinal univalence will also be lowered as bystanders may think their helping behavior is not urgently required in this case. This finding not only confirms the key role of social media platforms in bystander intervention but also supports the importance of person-environment interaction in explaining individuals' unethical or illegal behavior [25,54,55].

6.2. Implications for research

Our study has several theoretical implications. First, our work is one of the few studies investigating the factors influencing bystander inaction in delegitimizing doxing. While researchers across various fields have extensively examined cyberbullying in general [1,9,33], our study focuses on a unique form of cyberbullying, i.e., delegitimizing doxing, and thus responds to the calls [33,75] to differentiate forms of cyberbullying to avoid conceptual ambiguity and sharpen the research focus. Understanding why bystander inaction towards delegitimizing doxing is crucial because it is a potential way to cool down doxing incidents and keep the punishment of victims controllable and proportionate. Bystander inaction in delegitimizing doxing is not necessarily undesirable and this challenges the current understanding of bystander inaction that assumes it is a pro-bully behavior in cyberbullying [4,5,36]. Moreover, moving beyond prior studies of doxing, which have mostly examined the conceptual, moral, and legal issues of doxing [10,17] and the spreading mechanisms and consequences of doxing [30,32], our work develops an integrated model explaining bystander inaction in delegitimizing doxing. Therefore, our study not only advances the cyberbullying and doxing literature but also is expected to stimulate more academic discourses on this significant but underexplored phenomenon (i.e., delegitimizing doxing).

Second, to explain bystander inaction in delegitimizing doxing, we adopt interactionism (i.e., person-environment interaction) and combine attitudinal ambivalence theory and general deterrence theory to develop an integrated model. In contrast to established bystander intervention literature (which typically believes the causes of bystander inaction are diffusion of responsibility, evaluation apprehension, pluralistic ignorance, lack of empathy, lack of emergency awareness, and low self-efficacy [5,33,37]), our integrated model explains that bystander inaction is the result of the subjectivity of moral judgement and attitudinal ambivalence of bystanders and the roles of SNS in deterrence, i.e., the coexistence of positive and negative appraisals of delegitimizing doxing and the moderating effect of SNS deterrence. Specifically, we employed a congruence perspective to identify three distinctive congruence effects describing how different combinations of positive and negative appraisals determine bystander inaction in delegitimizing doxing. Our work also directs attention to the importance of person-environment interaction, i.e., the interplay of SNS deterrence and positive and negative attitudinal univalence shape bystander inaction, thus specifying the boundary condition within which attitudinal ambivalence determines bystander inaction. By delineating the distinctive antecedents of bystander inaction in delegitimizing doxing from previous studies, our work not only contributes to bystander intervention research by challenging the current understanding of bystander inaction and its causes in cyberbullying literature but also underscores the value of SNS deterrence to mitigate delegitimizing doxing.

Third, our study also contributes to attitudinal ambivalence literature. Prior studies have mostly considered attitudinal ambivalence as an overall state and measured it by using items directly expressing ambivalence [47] or by using a single index that arithmetically combines positive and negative evaluations [65]. Both methods have several methodological issues [18,57]. Drawing upon prior studies of congruence [69], we leverage PRRSA and uncover three novel influencing mechanisms that cannot be detected by prior studies, describing how different combinations of positive and negative appraisals relate to bystander inaction in delegitimizing doxing. As such, our study offers a more rigorous and innovative approach to test attitudinal ambivalence theory and thus advances attitudinal ambivalence research. Overall, our study provides theoretical and methodological contributions to attitudinal ambivalence literature.

6.3. Implications for practice

Our study also offers practical implications. First, bystander inaction in delegitimizing doxing results from the coexistence of both positive and negative appraisals of delegitimizing doxing. In particular, the

positive appraisal of delegitimizing doxing has a higher mean value than the negative appraisal of delegitimizing doxing in the data. Therefore, policymakers and platforms should distinguish between delegitimizing doxing and other forms of cyberbullying and take into account the uniqueness of delegitimizing doxing when formulating policies against cyberbullying. Specifically, policymakers and platforms should not completely block delegitimizing doxing but should recognize that doxing reflects netizens' pursuit of enforcement of social norms. This is actually one of the main reasons for bystanders to join in delegitimizing doxing, making delegitimizing doxing go viral online. To remedy this issue, policymakers and platforms should clearly define the boundaries of delegitimizing doxing and inform users of the devastating impact of delegitimizing doxing to make them fully and dialectically understand the phenomenon of delegitimizing doxing.

Moreover, perceived SNS deterrence weakens the negative impact of positive and negative attitudinal univalence on bystander inaction in delegitimizing doxing. That is, high levels of external sanctions on delegitimizing doxing on the platform will prevent bystanders from joining in delegitimizing doxing and helping the victims, indicating the likelihood of bystander inaction is less likely to be reduced. It suggests that social media platforms can play a proactive role in regulating delegitimizing doxing. To increase perceived SNS deterrence, platforms can implement robust detection and monitoring systems, issue clear countermeasures specific to delegitimizing doxing, and increase the enforcement of countermeasures against delegitimizing doxing to increase the certainty, severity, and celerity of sanctions against delegitimizing doxing.

6.4. Limitations and future research

The key findings and implications of our study should be interpreted acknowledging several limitations. First, we only integrated attitudinal ambivalence theory and general deterrence theory to explain bystander inaction in delegitimizing doxing. While we controlled several important factors (perceived empathy, personal responsibility, and demographics) in the model, we ignored other important factors, such as motivations for delegitimizing doxing, personality, social influences, and characteristics of delegitimizing doxing posts. We also didn't examine the determinants of positive and negative appraisals of delegitimizing doxing. Future research could follow this line to extend our model to provide a holistic view of bystander inaction in delegitimizing doxing. Second, while our study offers a novel and contextualized understanding of bystander inaction in delegitimizing doxing, we stop at the explanation of the phenomenon. We encourage future research to design and test the effectiveness of regulation strategies for delegitimizing doxing to provide policymakers and platforms with specific and actionable strategies. Third, we validated our model using cross-sectional survey data. To ensure response quality, we used several procedural and statistical remedies to detect and mitigate bias. We also asked respondents to answer survey questions based on a recalled delegitimizing doxing incident by themselves to reduce response bias. Due to the unique and complex nature of delegitimizing doxing, future research can design mixed-method research to provide richer details about the phenomenon of delegitimizing doxing.

7. Conclusion

Drawing upon attitudinal ambivalence theory and general deterrence theory, our study develops an integrative model to explain bystander inaction in delegitimizing doxing. We leverage PRRSA to test the effect of attitudinal ambivalence on bystander inaction in delegitimizing doxing, uncovering three influencing mechanisms: (1) bystander inaction remains unchanged as two equal levels of positive and negative appraisals increase in the same direction; (2) bystander inaction decreases as the degree of incongruence between positive and negative appraisals increases; and (3) positive attitudinal univalence has a stronger negative impact on bystander inaction, as compared to negative attitudinal univalence. We also find that perceived SNS deterrence weakens the negative impact of attitudinal univalence on bystander inaction. By providing a detailed and contextualized explanation of bystander inaction in the context of delegitimizing doxing, our study is expected to contribute to bystander intervention research and practice.

Funding

This work was substantially supported by grants from the National Natural Science Foundation of China (Project No 72301028, 72274144, 72311540158), the Humanities and Social Sciences Foundation of the Ministry of Education, China [22YJA870013], a fellowship award from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No HKBU SRFS2021-2H03), and the Fundamental Research Funds for the Central Universities, China.

CRediT authorship contribution statement

Yang-Jun Li: Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Zhi-Zhi He:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Christy M.K. Cheung:** Writing – review & editing, Writing – original draft, Supervision, Resources, Conceptualization. **Xiao-Liang Shen:** Writing – review & editing, Writing – original draft, Resources, Conceptualization. **Matthew K.O. Lee:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] P.B. Lowry, J. Zhang, C. Wang, M. Siponen, Why do adults engage in cyberbullying on social media? An integration of online disinhibition and deindividuation effects with the social structure and social learning model, Inf. Syst. Res. 27 (4) (2016) 962–986.
[2] J. Howarth, 17 cyberbullying facts & statistics (2024). https://explodingtopics.com/blog/cyberbullying-stats (2023).
[3] A. Fletcher, N. Fitzgerald-Yau, E.Allen R.Jones, R.M. Viner, C. Bonell, Brief report: cyberbullying perpetration and its associations with socio-demographics, aggressive behaviour at school, and mental health outcomes, J. Adolesc. 37 (8) (2014) 1393–1398.
[4] S.H. Taylor, D. DiFranzo, Y.H. Choi, S. Sannon, N.N. Bazarova, Accountability and empathy by design: encouraging bystander intervention to cyberbullying on social media, Proc. ACM Hum. -Comput. Interact. 3 (CSCW) (2019) 1–26.
[5] R.Y.M. Wong, C.M. Cheung, B. Xiao, J.B. Thatcher, Standing up or standing by: understanding bystanders' proactive reporting responses to social media harassment, Inf. Syst. Res. 32 (2) (2021) 561–581.
[6] L. You, Y.H. Lee, The bystander effect in cyberbullying on social network sites: anonymity, group size, and intervention intentions, Telemat. Inform. 45 (2019) 101284.
[7] Y. Seon, J.M. Swank, S. Smith-Adcock, A social interest–activity typology of bystander behavior in school bullying, J. Individ. Psychol. 77 (1) (2021) 24–39.
[8] D. DiFranzo, S.H. Taylor, F. Kazerooni, O.D. Wherry, N.N. Bazarova, Upstanding by design: bystander intervention in cyberbullying, in: Proc. SIGCHI Conf. Hum. Factor Comput. Syst, ACM, 2018, pp. 1–12.
[9] I. Coyne, A.M. Gopaul, M. Campbell, A. Pankász, R. Garland, F. Cousans, Bystander responses to bullying at work: the role of mode, type and relationship to target, J. Bus. Ethics. 157 (2019) 813–827.
[10] D.M. Douglas, Doxing: a conceptual analysis, Ethics Inf. Technol. 18 (3) (2016) 199–210.
[11] J.B. Grubbs, B. Warmke, J. Tosi, A.S. James, W.K. Campbell, Moral grandstanding in public discourse: status-seeking motives as a potential explanatory mechanism in predicting conflict, PLOS ONE 14 (10) (2019) e0223749.
[12] V. Ni, X. Zhu, China riveted by public row between pop star and former wife. https://www.theguardian.com/world/2021/dec/20/wang-leehom-lee-jinglei-china-taiwan, 2021.

[13] C. Capecchi, K. Rogers, Killer of Cecil the Lion finds out that he is a target now of internet vigilantism. https://www.nytimes.com/2015/07/30/us/cecil-the-lion-walter-palmer.html, 2015.
[14] C. Hatton, China's internet vigilantes and the "human flesh search engine". https://www.bbc.com/news/magazine-25913472, 2014.
[15] M.J. Crockett, Moral outrage in the digital age, Nat. Hum. Behav. 1 (11) (2017) 769–771.
[16] W.J. Brady, J.A. Wills, J.T. Jost, J.A. Tucker, J.J. Van Bavel, Emotion shapes the diffusion of moralized content in social networks, Proc. Natl. Acad. Sci. U.S.A. 114 (28) (2017) 7313–7318.
[17] J.M. MacAllister, The doxing dilemma: seeking a remedy for the malicious publication of personal information, Fordham Law Review 85 (5) (2017) 2451–2484.
[18] K. Jonas, P. Broemer, M. Diehl, Attitudinal ambivalence, Eur. Rev. Soc. Psychol. 11 (1) (2000) 35–74.
[19] P. Billingham, T. Parr, Enforcing social norms: the morality of public shaming, Eur. J. Philos. 28 (4) (2020) 997–1016.
[20] D.E. Leidner, O. Tona, The CARE theory of dignity amid personal data digitalization, MIS Q 45 (1) (2021) 343–370.
[21] B. Latane´, J.M. Darley, The Unresponsive bystander: Why doesn't He Help? Appleton-Century-Croft, New York, NY, 1970.
[22] M. Alvesson, J. Sandberg, Generating research questions through problematization, Acad. Manage. Rev. 36 (2) (2011) 247–271.
[23] S. Chatterjee, R.M. Davison, The need for compelling problematisation in research: the prevalence of the gap-spotting approach and its limitations, Inf. Syst. 31 (2) (2021) 227–230.
[24] J.P. Gibbs, Crime, punishment, and deterrence, SW Soc. Sci. Q. (1968) 515–530.
[25] Y.J. Li, C.M. Cheung, X.L. Shen, M.K. Lee, When socialization goes wrong: understanding the we-intention to participate in collective trolling in virtual communities, J. Assoc. Inf. Syst. 23 (3) (2022) 678–706.
[26] D. Trottier, Digital vigilantism as weaponisation of visibility, Philos Technol 30 (2017) 55–72.
[27] B. Anderson, M.A. Wood, Doxxing: a scoping review and typology, in: J. Bailey, A. Flynn, N. Henry (Eds.), The Emerald international Handbook of Technology Facilitated Violence and Abuse, Emerald Publishing Limited, Bingley, 2021, pp. 205–226.
[28] L.Y. Chang, R. Poon, Internet vigilantism: attitudes and experiences of university students toward cyber crowdsourcing in Hong Kong, Int. J. Offender Ther. Comp. Criminol. 61 (16) (2017) 1912–1932.
[29] G. Neubaum, L. Rösner, T. Ganster, K. Hambach, N.C. Krämer, United in the name of justice: how conformity processes in social media may influence online vigilantism, Psychol. Pop. Media Cult. 7 (2) (2018) 185–199.
[30] C.S. Heng, Z. Lin, X. Xu, Y. Zhang, Y. Zhao, Human flesh search: what did we find? Inf. Manag. 56 (4) (2019) 476–492.
[31] Q. Chen, K.L. Chan, A.S.Y. Cheung, Doxing victimization and emotional problems among secondary school students in Hong Kong, Int. J. Environ. Res. Public Health. 15 (12) (2018) 2665.
[32] I. Hansen, D.J. Lim, Doxing democracy: influencing elections via cyber voter interference, Contemp. Politics. 25 (2) (2019) 150–171.
[33] T.K. Chan, C.M. Cheung, Z.W. Lee, Cyberbullying on social networking sites: a literature review and future research directions, Inf. Manag. 58 (2) (2021) 103411.
[34] R. Thornberg, T. Jungert, Bystander behavior in bullying situations: basic moral sensitivity, moral disengagement and defender self-efficacy, J. Adolesc. 36 (3) (2013) 475–483.
[35] A. Sarmiento, M. Herrera-López, I. Zych, Is cyberbullying a group process? Online and offline bystanders of cyberbullying act as defenders, reinforcers and outsiders, Comput. Hum. Behav. 99 (2019) 328–334.
[36] T.K. Chan, C.M. Cheung, I. Benbasat, B. Xiao, Z.W. Lee, Bystanders join in cyberbullying on social networking sites: the deindividuation and moral disengagement perspectives, Inf. Syst. Res. 34 (3) (2023) 828–846.
[37] P. Fischer, J.I. Krueger, T. Greitemeyer, C. Vogrincic, A. Kastenmüller, D. Frey, M. Kainbacher, The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies, Psychol. Bull. 137 (4) (2011) 517–537.
[38] R. Moerland, Individuals As Bystanders to Atrocity crimes, in: The Oxford Handbook of Atrocity Crimes, Oxford University Press, 2022, pp. 303–328.
[39] I. Ajzen, M. Fishbein, Attitudes and the attitude-behavior relation: reasoned and automatic processes, Eur. Rev. Soc. Psychol. 11 (1) (2000) 1–33.
[40] B.E. Ashforth, K.M. Rogers, M.G. Pratt, C. Pradies, Ambivalence in organizations: a multilevel approach, Organ. Sci. 25 (5) (2014) 1453–1478.
[41] A. Lakhiwal, H. Bala, P.M. Léger, Ambivalence is better than indifference: behavioral and neurophysiological assessment of ambivalence in online environments, MIS Q 47 (2) (2023) 705–732.
[42] X.L. Shen, Y.J. Li, Y. Sun, Wearable health information systems intermittent discontinuance: a revised expectation-disconfirmation model, Ind. Manag. Data Syst. 118 (3) (2018) 506–523.
[43] K.C. Ng, X. Zhang, J.Y. Thong, K.Y. Tam, Protecting against threats to information security: an attitudinal ambivalence perspective, J. Manag. Inf. Syst. 38 (3) (2021) 732–764.
[44] C. Dubreuil, D. Dion, S. Borraz, For the love of the game: moral ambivalence and justification work in consuming violence, J. Bus. Ethics. 186 (3) (2023) 675–694.
[45] B. Wu, S.C. Schuh, H. Wei, Y. Cai, When positives and negatives collide: evidence for a systematic model of employees' strategies for coping with ambivalence, J. Bus. Psychol. 38 (2) (2023) 473–491.
[46] N.B. Rothman, M.G. Pratt, L. Rees, T.J. Vogus, Understanding the dual nature of ambivalence: why and when ambivalence leads to good and bad outcomes, Acad. Manag. Ann. 11 (1) (2017) 33–72.
[47] H. Qahri-Saremi, O. Turel, Ambivalence and coping responses in post-adoptive information systems use, J. Manag. Inf. Syst. 37 (3) (2020) 820–848.
[48] X.L. Jin, Z. Zhou, Y. Tian, A configurational analysis of the causes of consumer indirect misbehaviors in access-based consumption, J. Bus. Ethics. 175 (2022) 135–166.
[49] J. Farrar, T. King, To punish or not to punish? The impact of tax fraud punishment on observers' tax compliance, J. Bus. Ethics. 183 (1) (2023) 289–311.
[50] A.C. Johnston, M. Warkentin, M. Siponen, An enhanced fear appeal rhetorical framework, MIS Q 39 (1) (2015) 113–134.
[51] P.B. Lowry, C. Posey, R.B.J. Bennett, T.L. Roberts, Leveraging fairness and reactance theories to deter reactive computer abuse following enhanced organisational information security policies: an empirical study of the influence of counterfactual reasoning and organisational trust, Inf. Syst. 25 (3) (2015) 193–273.
[52] B. Xu, Z. Xu, D. Li, Internet aggression in online communities: a contemporary deterrence perspective, Inf. Syst. 26 (6) (2016) 641–667.
[53] K.Y. Tam, K.Y. Feng, S. Kwan, The role of morality in digital piracy: understanding the deterrent and motivational effects of moral reasoning in different piracy contexts, J. Assoc. Inf. Syst. 20 (5) (2019) 604–628.
[54] R. Willison, M. Warkentin, A.C. Johnston, Examining employee computer abuse intentions: insights from justice, deterrence and neutralization perspectives, Inf. Syst. 28 (2) (2018) 266–293.
[55] B. Guan, C. Hsu, The role of abusive supervision and organizational commitment on employees' information security policy noncompliance intention, Internet Res 30 (5) (2020) 1383–1405.
[56] L.R. Shanock, B.E. Baran, W.A. Gentry, S.C. Pattison, E.D. Heggestad, Polynomial regression with response surface analysis: a powerful approach for examining moderation and overcoming limitations of difference scores, J. Bus. Psychol. 25 (2010) 543–554.
[57] V. Venkatesh, S. Goyal, Expectation disconfirmation and technology adoption: polynomial modeling and response surface analysis, MIS Q 34 (2) (2010) 281–303.
[58] M.K. Stein, S. Newell, E.L. Wagner, R.D. Galliers, Coping with information technology: mixed emotions, vacillation, and nonconforming use patterns, MIS Q 39 (2) (2015) 367–392.
[59] Y.A. Yao, Z. Ma, Toward a holistic perspective of congruence research with the polynomial regression model, J. Appl. Psychol. 108 (3) (2023) 446–465.
[60] X.L. Shen, Y.J. Li, Y. Sun, F. Wang, Good for use, but better for choice: a relative model of competing social networking services, Inf. Manag. 58 (3) (2021) 103448.
[61] H.T. Yi, M. Lee, F.E. Amenuvor, Effects of contractual completeness on opportunistic behaviors, multilevel transaction costs and relationship termination intention: the moderating effect of alternative attractiveness, Eur. J. Mark. 57 (11) (2023) 2945–2973.
[62] Z.R. Steelman, B.I. Hammer, M. Limayem, Data collection in the digital age: innovative alternatives to student samples, MIS Q 38 (2) (2014) 355–378.
[63] H. Aguinis, I. Villamor, R.S. Ramani, MTurk research: review and recommendations, J Manage 47 (4) (2021) 823–837.
[64] G. Klein, J.J. Jiang, P. Cheney, Resolving difference score issues in information systems research, MIS Q 33 (4) (2009) 811–826.
[65] Y. Cornil, N. Ordabayeva, U. Kaiser, B. Weber, P. Chandon, The acuity of vice: attitude ambivalence improves visual sensitivity to increasing portion sizes, J. Consum. Psychol. 24 (2) (2014) 177–187.
[66] G.L. Polites, N. Roberts, J. Thatcher, Conceptualizing models using multidimensional constructs: a review and guidelines for their use, Eur. J. Inf. Syst. 21 (2012) 22–48.
[67] P.M. Podsakoff, S.B. MacKenzie, J.Y. Lee, N.P. Podsakoff, Common method biases in behavioral research: a critical review of the literature and recommended remedies, J. Appl. Psychol. 88 (5) (2003) 879–903.
[68] S.J. Breckler, A comparison of numerical indexes for measuring attitude ambivalence, Educ. Psychol. Meas. 54 (2) (1994) 350–365.
[69] J.R. Edwards, Alternatives to difference scores: polynomial regression and response surface methodology, in: J. Drasgow, N. Schmitt (Eds.), Measuring Analyzing Behavior organizations: Advances in Measurement and Data Analysis, Jossey-Bass Inc, 2002, pp. 350–400.
[70] C.Y. Tsai, J. Kim, F. Jin, M. Jun, M. Cheong, F.J. Yammarino, Polynomial regression analysis and response surface methodology in leadership research, Leadersh. Q. 33 (1) (2022) 101592.
[71] A. Benlian, Effect mechanisms of perceptual congruence between information systems professionals and users on satisfaction with service, J. Manag. Inf. Syst. (2013) 63–96.
[72] J.R. Edwards, M.E. Parry, On the use of polynomial regression equations as an alternative to difference scores in organizational research, Acad. Manag. J. 36 (6) (1993) 1577–1613.
[73] D.C. Chau, E.W. Ngai, J.E. Gerow, J.B. Thatcher, The effects of business-IT strategic alignment and IT governance on firm performance: a moderated polynomial regression analysis, MIS Q 44 (4) (2020) 1679–1703.
[74] Y. Sun, Y. Fang, K.H. Lim, D. Straub, User satisfaction with information technology service delivery: a social capital perspective, Inf. Syst. Res. 23 (4) (2012) 1195–1211.
[75] H. Scheithauer, A. Schultze-Krumbholz, J. Pfetsch, M. Hess, Types of cyberbullying, in: P.K. Smith, J.O.H. Norman (Eds.), The Wiley Blackwell Handbook of bullying: A comprehensive and International Review of Research and Intervention, Wiley Blackwell, 2021, pp. 120–138.
[76] Meta. Facebook community standards. https://transparency.meta.com/zh-cn/policies/improving/proactive-rate-metric/(2024).

[77] H.-T. Tsai, R.P. Bagozzi, Contribution behavior in virtual communities: cognitive, emotional, and social influences, MIS Q. 38 (1) (2014) 143–164.
[78] J.-B.E.M. Steenkamp, A. Maydeu-Olivares, An updated paradigm for evaluating measurement invariance incorporating common method variance and its assessment, J. Acad. Market Sci. 49 (1) (2021) 5–29.
[79] A.E. Smith, N. Zlatevska, R.M.M.I. Chowdhury, A. Belli, A meta-analytical assessment of the effect of deontological evaluations and teleological evaluations on ethical judgments/intentions, J. Bus. Ethics. 188 (3) (2023) 553–588.

Yang-Jun Li is an assistant professor in Department of Management Engineering at Beijing Institute of Technology. His research interests include ethical issues in the use and application of IT, e-commerce, knowledge management, and social media. He has published in journals such as *Journal of the Association for Information Systems, Information & Management, Decision Support Systems, Journal of the Association for Information Science and Technology,* and *Journal of Business Research*, among others.

Zhi-Zhi He is a PhD candidate in the Department of Information Systems at City University of Hong Kong. Her research interests include digitalization, social media, and societal implications of technology use. Her works have appeared in the proceedings of International Conference on Information Systems and Pacific Asia Conference on Information Systems.

Christy M. K. Cheung is a Chair professor in Information Systems and Digital Innovation Management at Hong Kong Baptist University. Her research interests includee IT and user behaviors, social media and electronic commerce, and IT and societal implications. Her work appears in *Information Systems Research, Journal of the Association for Information Systems, Information & Management, Journal of Management Information Systems,* and *MIS Quarterly*, and many others. She serves as the editor-in-chief at Internet Research.

Xiao-Liang Shen is a professor in the School of Information Management at Wuhan University. His current research interests include information management, the dark side of IT, and digital governance. He has published in journals such as *Journal of the Association for Information Systems, Journal of Information Technology, Journal of the Association for Information Science and Technology, Information & Management,* and *Decision Support Systems*.

Matthew K. O. Lee is a chair professor of information systems and e-commerce in the Department of Information Systems at City University of Hong Kong. His research interests include IT-based innovation adoption and diffusion, knowledge management, electronic commerce, online addiction, and the development of digital competence. Professor Lee has published over one hundred refereed articles in international journals, such as *MIS Quarterly, Journal of MIS, Journal of AIS, Information & Management,* and *Journal of International Business Studies*, among others. He has been rated as the most highly cited business professor at CityU for many years. Professor Lee has taken an editorial role in a number of leading scholarly journals including *MIS Quarterly, Information Systems Journal*, and *International Journal of Information Management*, among others.