



Celebrity worship and online trolling in college students: Self-esteem and moral disengagement mediation and perceived anonymity moderation

Mengni Yu¹, Wenyu Zeng¹, Yihong Yang¹, Biyun Wu² and Fang Li^{1,3,*}

¹School of Psychology, Sichuan Normal University, Chengdu, China

²The Third Clinical Medical College of Guangzhou University of Chinese Medicine, Guangzhou, China

³Sichuan Key Laboratory of Psychology and Behaviour of Discipline Inspection and Supervision, Sichuan Normal University, Chengdu, China

*Correspondence: Fang Li, lifang_sicnu@126.com

Received: 19 May 2025; Accepted: 19 May 2026; Published: 31 August 2026

Abstract: We examined whether celebrity worship predicts trolling via self-esteem and moral disengagement, and whether perceived anonymity moderates these effects. A total of 1057 college students (females = 54.6%; mean age = 20.09 years, SD = 1.734) completed the measurements of celebrity worship, self-esteem, moral disengagement, perceived anonymity, and online trolling. Structural equation modeling (SEM) results showed that higher celebrity worship was associated with online trolling. Lower self-esteem and moral disengagement respectively played mediating roles between them, exacerbating trolling behavior. Self-esteem and moral disengagement also played a sequential mediating role for higher trolling behavior. Perceived anonymity moderated both the direct and indirect paths between celebrity worship and online trolling to be stronger. With a high level of perceived anonymity, the predictive effects of celebrity worship and moral disengagement on online trolling were exacerbated. Consistent with the I^3 model, celebrity worship and moral disengagement enhance online trolling tendencies, and such effects are amplified when people perceived to be anonymous. Those of self-esteem may be less susceptible to these tendencies. Interventions aimed at reducing college students' online trolling risk from celebrity worship should include strategies to rebuild self-concept and clarify speech boundaries for rational discourse this digital age.

Keywords: celebrity worship; online trolling; self-esteem; moral disengagement; perceived anonymity

Introduction

Celebrities often attract a large and sometimes intense following among college students, especially the younger individuals (Giles & Maltby, 2004; Ouvrein et al., 2018). Many are inclined to follow celebrities online on an ongoing basis for vicarious gratification to an extent that incline towards worshipping. Celebrity worship refers to the fact that an individual shows their attention and emotional projection towards the person they subjectively recognize and find inspiring. It encompasses both the individual's internal identification with and attachment to this person, as well as their external behavioral tendencies (Li et al., 2023). In their celebrity worship interactions, some may engage in online trolling and inadvertently share opinions that may provoke others or escalate into arguments (Nitschinsk et al., 2022). The likelihood for conflicts from celebrity-related discussions may be shaped by self-esteem statuses and moral disengagement predispositions in ways yet to be reported on. Self-esteem refers how a person views and evaluates themselves, and it plays a crucial role in shaping their overall personality. Moral disengagement is when individuals less sensitive to hurtfulness to others and the suffering endured by the victim (Bandura, 1999).

Celebrity worship and online trolling. Currently, the phenomenon of celebrity worship in college students' culture is becoming increasingly prevalent and has even become a part of the daily lives of most college students (Cui et al., 2022). Through celebrity worship, they can fulfill their needs for belonging, emotional catharsis, and social identification during the exploration of self-identity

(Pirzade et al., 2024). On the one hand rational celebrity worship energize people serving as a driving force for striving, and promoting their self-growth and progress. On the other hand, irrational celebrity workshop may trigger intense online confrontations that could escalate into group conflicts, including cyberbullying (Zhang et al., 2025).

Individual thought patterns and group beliefs. Firstly, at the individual cognitive level, excessive celebrity worship will weaken an individual's cognitive flexibility and form a black-and-white thinking pattern (Maltby et al., 2004), which refers to a rigid and absolutist cognitive style that simplistically classifies things into two opposing extremes while disregarding potential intermediate states. Moreover, individuals may have an information bias that could be amplified in the "echo chamber" environment constituted by algorithmic recommendations and homogeneous communities (Putri et al., 2024). Secondly, in online communities, netizens (i.e., Individuals who actively participate in, contribute to, and are committed to improving online communities; Hauben & Hauben, 1997) may have a bias towards certain information (Putri et al., 2024), and individual self-awareness is submerged by group identity, leading to a state of 'deindividuation'. This means that individuals within a group lose their sense of self and decisional autonomy being subservient to group norms in their personal rational judgment. Furthermore, the disinhibition effect of online anonymity could make people more prone to engaging in aggressive acts on the internet. Therefore, we speculate that higher celebrity worship is associated with online trolling behaviour.



Self-esteem mediation. Empirical studies have reported lower online trolling with high self-esteem (Zezulka & Seigfried-Spellar, 2016). For instance, individuals with low self-esteem tend to ward off feelings of inferiority by shifting blame to others through hostility and aggression (Tracy & Robins, 2003), any may be predisposed to cyberbullying or other harmful acts. With higher celebrity worship increases, their self-esteem levels tend to decline (Stepanyan, 2019), perhaps from upward social comparisons in which they self perceive as less worthy (Fox & Vendemia, 2016). Therefore, we speculate that self-esteem mediates the relationship between celebrity worship and online trolling.

Moral disengagement mediation. Moral disengagement is a stable predictor of aggressive behavior (Runions & Bak, 2015), characterized by euphemistic labeling, and irrational advantageous self-comparison (Bandura, 1999). Thus, moral disengagement serves as a crucial psychological bridge connecting online self-entertainment, cognitive rationalization, and conflict behavior (Saulnier & Krettenauer, 2023). Even if the perpetrator recognizes the immorality of their actions, moral disengagement effectively inhibit the self-condemnation arising during the commission of aggressive acts, while alleviating the psychological discomfort triggered by the violation of moral principles (Gasser & Keller, 2009). Apparently, celebrity worship is associated with moral disengagement, with the depth of worship being directly proportional to the degree of disengagement and desire for in-group favoritism rewards (Ouvrein et al., 2018; Tajfel et al., 1979). Thus, we propose that moral disengagement mediates the association between celebrity worship and online trolling.

Perceived anonymity as a moderator. A key feature of online is its anonymity (Wright, 2013) and disinhibition provided by digital platforms (Singh, 2024). When individuals perceive themselves as anonymous online, they may be susceptible to online trolling behaviors (Wachs & Wright, 2018). Specifically, anonymity removed the social constraints and for real-world consequences. It can also strengthen the norms of the immediate situation, reduce accountability, and increase the likelihood of provocative, abusive, and other uncivil behaviors on social media platforms (Sude & Dvir-Gvirsman, 2023).

Self-esteem–moral disengagement chain mediation with perceived anonymity. As previously noted, low self-esteem tend to exhibit a stronger tendency towards aggression under anonymous conditions (Thomaes et al., 2008), and would be predisposed to online trolling as a self-worth compensation strategy in an anonymous environment (Sest & March, 2017). These effects are amplified with diffusion of responsibility in moral disengagement to a group identity and weakening the perception of behavioral consequences. Consequently, perceived anonymity plays a moderating role between the influencing factors of online aggressive behavior and online aggressive behavior itself (Chu et al., 2020). Since online trolling is one of the current new types of online aggressive behavior, we speculate that perceived anonymity can moderate both the direct and indirect paths between celebrity worship and online trolling. Specifically, perceived anonymity can enhance the promoting effects of celebrity worship and

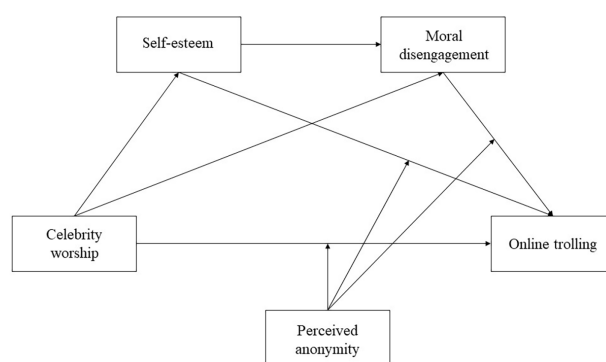


Figure 1. Hypothesized model diagram

moral disengagement on online trolling, and weaken the mitigating effect of self-esteem on online trolling.

Theoretical basis. The I³ model proposes that an individual's aggressive behavior is influenced by three aspects: instigation, impellance, and inhibition. Instigation refers to situational factors or relatively stable factors. Impellance refers to the clues or situations that induce an individual to take action against a target object. Both of these are promoting forces for an individual's aggressive behavior. Inhibition, however, is a factor that helps an individual reduce the tendency of aggressive behavior. It is a negative force that weakens the occurrence of an individual's aggressive behavior (Finkel, 2014). During the occurrence and development of an individual's behavior, these three elements can act independently, or they can interact with and influence one another, jointly contributing to the generation of the behavior (Finkel & Hall, 2018).

Based on the dynamic interaction logic of the I³ model, in the specific context of online trolling, celebrity worship acts as a stable internal impellance, sustaining a baseline aggressive tendency and motivating protective behaviors for the idol or group. Moral disengagement serves as a cognitive impellance, reducing moral constraints and priming individuals for aggression. Self-esteem functions as the key inhibition, with higher levels bolstering resistance to identity-based or morally disengaged impulses.

The present study

We investigated how celebrity worship relates to online trolling behavior, with a focus on the mediating effects of self-esteem and moral disengagement and the moderating effect of perceived anonymity (see Figure 1). Specifically, the following five hypotheses are proposed:

Hypothesis 1: Higher celebrity worship is associated with risk for addictive online trolling.

Hypothesis 2: Lower self-esteem mediates the relationship between celebrity worship and online trolling for higher online trolling.

Hypothesis 3: Moral disengagement mediates the relationship between celebrity worship and online trolling for higher online trolling.

Hypothesis 4: Self-esteem and moral disengagement play a sequential mediation role between celebrity worship and online trolling, for higher online trolling.

Hypothesis 5: Perceived anonymity plays a moderating role in both the direct and indirect paths between celebrity worship and online trolling, such that these effects are stronger when perceived anonymity is higher.

Method

Participants and setting

Participants were 1057 students from three universities in Chengdu, China. They were predominantly trained post-graduate students majoring in psychology. Among them, there were 480 males (45.4%) and 577 females (54.6%). The average age of the research subjects was 20.09 ± 1.734 years old.

Measures

Celebrity worship was operationalized through **the Chinese version of the Celebrity Attitude Scale** (Peng et al., 2010). This 27-item tool uses a 5-point Likert response format, and total scores correspond to worship intensity (higher scores = greater celebrity worship). In this research, Cronbach's α was 0.962.

The Self-esteem Scale (Rosenberg, 1965) was administered to capture participants' self-esteem. This 10-item measure employs a 4-point Likert format; relevant items were reverse-coded, with total scores reflecting the construct. In this research, Cronbach's α was 0.855.

The Moral Disengagement Scale (Yang & Wang, 2011; adapted from Bandura et al., 1996) was administered. This 32-item tool employs a 5-point Likert format, with total scores indexing moral disengagement. In this research, Cronbach's α was 0.944.

For the measurement of perceived anonymity, this study adopted **the Perceived Anonymity Scale** developed by Jung et al. (2012) for assessment. This scale uses a 7-point Likert scale and consists of 4 items. The sum of the scores of the 4 items is the perceived anonymity of the participants. In this research, Cronbach's α was 0.828.

Online trolling was assessed with **the Chinese version of the Online Trolling Scale** (Sest & March, 2017; translated by Li et al., 2024). This 8-item measure uses a 5-point Likert format, with total scores indexing the behavior's frequency. Cronbach's α in the current sample was 0.815.

Procedure

The Ethics Review Committee of Sichuan Normal University granted ethical approval (IRB No. SCNU-230904). Consenting participants were assured that their data were solely for academic research and statistical analysis, will not be disclosed to any third party under any circumstances, and they may terminate their participation at any time during the questionnaire completion process. They elected to take paper or online versions of the surveys. For paper questionnaires, sequential numbers were used to replace personal identifiers during data entry. Furthermore, online questionnaires were distributed via a professional survey platform, where functions such as IP address tracking and device information logging were disabled. Although offline questionnaires were distributed with on-site instructions, responses were still completed via QR-code mobile submission, making online and offline response modes essentially identical.

Data analysis

Descriptive statistics, independent samples *t*-test, mediation and moderation effect analyses, and multi-group analysis, among others, were conducted on the data. First, we employed a single-factor confirmatory factor analysis to test for common method bias by constructing a model in which all items were loaded onto a single latent variable. If the single-factor model had demonstrated good fit, it would have indicated severe common method bias. The results showed that the model fit was extremely poor: $\chi^2/df = 10.238 > 5$, CFI = $0.441 < 0.9$, TLI = $0.426 < 0.9$, RMSEA = $0.093 > 0.08$, SRMR = $0.126 > 0.08$. These outcomes demonstrate that common method bias is not a major problem in the present study.

Since gender, age, and residential address showed significant correlations with the main research variables, these variables were included as controls. Celebrity worship was treated as the independent variable, self-esteem and moral disengagement as mediators, and online trolling as the dependent variable. A structural equation model was constructed to analyze the chain mediating effects of self-esteem and moral disengagement between celebrity worship and online trolling.

In the supplementary analysis, using structural equation modeling (SEM), we conducted a multi-group analysis by gender based on the chain mediation model and compared the moderated mediation models across genders to examine potential gender differences. Based on the mediation analysis, we further examined the moderating effect of perceived anonymity.

Results

Descriptive statistical results

The results of the Pearson correlation analysis are presented in Table 1. Celebrity worship was also positively correlated with moral disengagement and perceived anonymity ($r = 0.378, 0.258, ps < 0.001$), while it was negatively correlated with self-esteem ($r = -0.112, p < 0.001$). Self-esteem was negatively correlated with moral disengagement, perceived anonymity, and online trolling ($r = -0.370, -0.201, -0.326, ps < 0.001$). To further explore the association between celebrity worship and online trolling, this study calculated the mean score for each of the five dimensions of the celebrity worship scale—entertainment-social, emotional projection, complete identification, relational fantasy, and pathological edge—and conducted Pearson correlation analyses between these dimension scores and online trolling scores. The results indicated that all dimensions of celebrity worship were significantly positively correlated with online trolling ($ps < 0.001$), with variations in the strength of the correlations. Overall, the findings exhibited a characteristic pattern wherein higher levels of worship involvement were associated with stronger correlations with online trolling. Specifically, the entertainment-social dimension (low involvement) demonstrated the weakest positive correlation with online trolling ($r = 0.228, p < 0.001$), classified as a weak positive correlation; the emotional projection dimension ($r = 0.332, p < 0.001$) and relational fantasy dimension ($r = 0.335, p < 0.001$) showed similar

Table 1. Results of correlation analysis among variables

Variable	M ± SD	1	2	3	4	5	6	7	8
1. Gender	—	—							
2. Age	20.090 ± 1.734	0.001	—						
3. Residential address	—	-0.019	-0.030	—					
4. Celebrity worship	2.379 ± 0.925	0.072*	0.293***	-0.094**	—				
5. Self-esteem	2.819 ± 0.540	0.128***	0.029	-0.032	-0.112***	—			
6. Moral disengagement	1.977 ± 0.651	-0.207***	0.074*	-0.065*	0.378***	-0.370***	—		
7. Perceived anonymity	4.308 ± 1.493	-0.009	0.036	-0.036	0.258***	-0.201***	0.285***	—	
8. Online trolling	1.589 ± 0.657	-0.174***	0.164***	-0.091**	0.391***	-0.326***	0.609***	0.190***	—

Note. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 2. Bootstrap test for mediating effects

Model paths	Mediation effect	95% LLCI	95% ULCI
1. Celebrity worship → Self-esteem → Online trolling	0.015	0.007	0.026
2. Celebrity worship → Moral disengagement → Online trolling	0.165	0.127	0.212
3. Celebrity worship → Self-esteem → Moral disengagement → Online trolling	0.018	0.009	0.029
4. Total effect	0.368	0.309	0.427

correlation strengths with online trolling, both categorized as moderate-weak positive correlations; the pathological edge dimension (high involvement) was moderately positively correlated with online trolling ($r = 0.448$, $p < 0.001$); and the complete identification dimension (moderate-high involvement) yielded the strongest correlation with online trolling ($r = 0.481$, $p < 0.001$), representing a moderate positive correlation.

Celebrity worship and trolling behavior

The regression analysis with celebrity worship as the independent variable and online trolling as the dependent variable revealed that celebrity worship had a significant positive effect on online trolling ($b = 0.368$, $p < 0.001$). Specifically, celebrity worship predicted an increase in the level of online trolling, thereby validating Hypothesis 1.

Analysis of the mediating effects of self-esteem and moral disengagement

The structural equation model demonstrated a good fit ($\chi^2/df = 6.943$, RMSEA = 0.075, GFI = 0.964). The results showed that celebrity worship positively predicted online trolling ($b = 0.170$, $p < 0.001$), negatively predicted self-esteem ($b = -0.122$, $p < 0.001$), and positively predicted moral disengagement ($b = 0.353$, $p < 0.001$); self-esteem negatively predicted moral disengagement ($b = -0.307$, $p < 0.001$) and online trolling ($b = -0.119$, $p < 0.001$); and moral disengagement positively predicted online trolling ($b = 0.468$, $p < 0.001$). The results of the significance test of the mediation tests showed that (see Table 2) the mediating path of celebrity worship → self-esteem → online trolling was significant ($b = 0.015$, 95% CI = [0.007, 0.026]), indicating that self-esteem mediated the relationship between celebrity worship and online trolling, which supported Hypothesis 2; the mediating path of celebrity worship → moral disengagement → online trolling was significant ($b = 0.165$, 95% CI = [0.127, 0.212]), indicating that moral disengagement mediated the relationship between celebrity worship and online trolling, verifying Hypothesis 3; the mediating

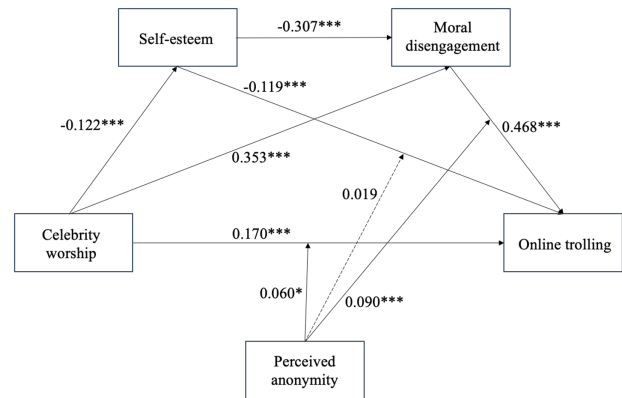


Figure 2. A moderated mediation model diagram. Note. * $p < 0.05$, *** $p < 0.001$.

path of celebrity worship → self-esteem → moral disengagement → online trolling was significant ($b = 0.018$, 95% CI = [0.009, 0.029]), indicating that self-esteem and moral disengagement had a chained mediating effect between celebrity worship and online trolling, supporting Hypothesis 4.

Analysis of the moderating effect of perceived anonymity

Figure 2 shows the moderated mediation model diagram. The results showed that the interaction term between celebrity worship and perceived anonymity has a significant impact on online trolling ($b = 0.060$, $p < 0.05$), indicating that the moderating effect of perceived anonymity between celebrity worship and online trolling was significant; the interaction term of moral disengagement and perceived anonymity had a significant effect on online trolling ($b = 0.090$, $p < 0.001$), indicating that the moderating effect of perceived anonymity between moral disengagement and online trolling was significant, partially verifying Hypothesis 5.

A simple slope test was further conducted. As shown in Figure 3, among individuals with high perceived anonymity ($b_{simple} = 0.230$, $p < 0.001$), the positive

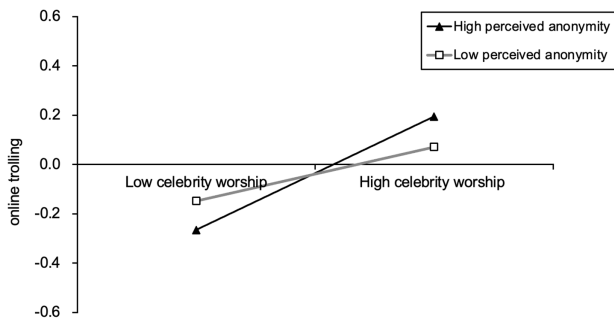


Figure 3. The relationship between celebrity worship and online trolling among individuals at different levels of perceived anonymity

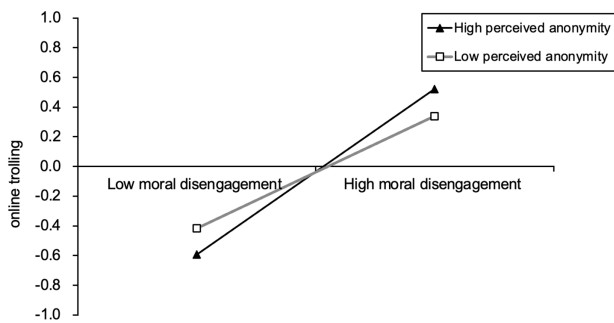


Figure 4. The relationship between moral disengagement and online trolling among individuals at different levels of perceived anonymity

predictive effect of celebrity worship on online trolling was higher than that among individuals with low perceived anonymity ($b_{\text{simple}} = 0.110, p < 0.01$). This indicates that perceived anonymity strengthens the effect of celebrity worship on young people's online trolling. As shown in Figure 4, among individuals with high perceived anonymity ($b_{\text{simple}} = 0.558, p < 0.001$), the positive predictive effect of moral disengagement on online trolling was higher than that among individuals with low perceived anonymity ($b_{\text{simple}} = 0.378, p < 0.001$). However, the research results found that the interaction term of self-esteem and perceived anonymity had no significant effect on online trolling ($b = 0.019, p > 0.05$), indicating that the moderating effect of perceived anonymity between self-esteem and online trolling was not significant.

Collateral findings

First, in the multi-group analysis, a freely estimated model (M1) and a constrained model (M2) were constructed respectively, and nested model comparison was conducted. The results showed that both M1 and M2 had good model fits, and there was a significant difference between the two models ($\Delta\chi^2/df = 5.248, p < 0.001$), indicating that the serial mediating effect exhibited a significant gender difference. Further comparisons of the specific path coefficients in the structural equation models for males and females are presented. The results indicated that three path coefficients—celebrity worship $\rightarrow$ online trolling ($p < 0.01$), celebrity worship $\rightarrow$ self-esteem ($p < 0.05$), and self-esteem $\rightarrow$ moral disengagement ($p < 0.05$)—exhibited significant gender differences, while the other

path coefficients did not show significant differences ($p > 0.05$). Specifically, compared with females, celebrity worship had a stronger predictive effect on self-esteem and online trolling among males, and self-esteem had a stronger predictive effect on moral disengagement among males.

Second, gender-comparative analyses of the moderated mediation model yielded distinct results. Among male participants, significant indirect effects were observed for moral disengagement and for the self-esteem $\rightarrow$ moral disengagement chain, whereas self-esteem alone failed to mediate. Perceived anonymity moderated only the path from moral disengagement to online trolling. Among female participants, all three mediating pathways attained significance; perceived anonymity again exerted moderation exclusively on the moral disengagement–trolling path. Detailed outputs appear in Supplementary Tables S1–S2 and Figures S1–S4.

Discussion

Analyses showed that celebrity worship was significantly and positively correlated with online trolling in the university student sample. Due to the convenience of the Internet, individuals engage with like-minded netizens on platforms such as INS, Weibo, and WeChat, fostering a sense of belonging and identity (Pirzade et al., 2024). Social identity theory posits that in-group favoritism emerges among members of the same group, wherein they exhibit more positive evaluations of their own group and negative biases toward out-groups (Tajfel et al., 1979). Furthermore, inter-group comparisons during such interactions often trigger group polarization—a phenomenon where specific viewpoints within the group are amplified and radicalized (Van Swol, 2009). Heightened levels of celebrity worship contribute to cognitive rigidity and selective information processing, which psychologically legitimizes aggressive behaviors, ultimately escalating online trolling among college students.

Celebrity worship exhibited a significant positive association with online trolling behavior, aligning with previous evidence linking celebrity worship to cyberbullying (Zhang et al., 2025). When celebrity worship mainly remains at the entertainment-social level, it more satisfies individual social needs and entertainment needs rather than deep emotional dependence (McCutcheon et al., 2002). This may be because individuals begin to project emotional needs onto celebrities and may develop virtual relational fantasies (Li et al., 2023). In contrast, the strong correlations between the complete identification dimension and pathological edge dimension with online trolling further confirm the danger of extreme worship forms (McCutcheon et al., 2002). When individuals completely internalize celebrities as part of their self-concept, maintaining the celebrity's image becomes equivalent to maintaining self-worth (Li et al., 2023).

The mediating effect of self-esteem between celebrity worship and online trolling was significant. Specifically, as individuals' celebrity worship intensifies, they may develop pathological fixation (McCutcheon et al., 2002), and by negative self comparisons between themselves and idols (Tesser et al., 1988). Such comparisons often result

in frustration and self-deprecation, thereby lowering self-esteem (He & Xu, 2023). Low self-esteem is characterized by reduced confidence and impaired emotional regulation, which compromises rational responding during stressful situations (Tracy & Robins, 2003). Consequently, they are more prone to online trolling behaviors.

The mediating effect of moral disengagement between celebrity worship and online trolling was significant. This may stem from the online communities where college students with high celebrity worship are embedded (Pirzade et al., 2024). When engaging in conflict-prone behaviors like online trolling, individuals may deflect personal responsibility by emphasizing their group membership (Postmes & Spears, 1998), thereby distributing blame across the collective, which is also consistent with moral disengagement theory (Bandura, 1999). The activation of moral disengagement mechanisms allows them to act more recklessly, as perceived shared responsibility mitigates negative emotions directly triggered by aggression (Runions & Bak, 2015). Consequently, reduced guilt or anxiety further amplifies their likelihood of initiating online trolling.

Celebrity worship exerted a significant indirect effect on online trolling through the sequential mediators of self-esteem and moral disengagement. This could be explained by low self-esteem and celebrity worship intensifying upward social comparisons result in a progressive decline in self-esteem. To mitigate this cognitive dissonance, individuals may activate moral disengagement mechanisms. These mechanisms buffer guilt associated with immoral behaviors, ultimately facilitating online trolling.

Perceived anonymity significantly moderated the relationships of both celebrity worship and moral disengagement with online trolling. First, higher perceived anonymity strengthened the predictive effect of celebrity worship on online trolling. This may occur because anonymity decouples online identity from offline social roles, reducing social constraints and accountability concerns (Suler, 2004). Moreover, online disinhibition, lowers the psychological costs of aggression (e.g., social evaluation anxiety), and helps translate sacralized cognitions about the celebrity/stance into trolling motives (Suler, 2004; Postmes & Spears, 1998). Second, the positive association between moral disengagement and online trolling intensified with elevated perceived anonymity. High anonymity may further reduce perceived accountability and concerns about offline consequences, thereby increasing the likelihood of misconduct (Sude & Dvir-Gvirsman, 2023). Concurrently, moral disengagement (e.g., framing trolling as “upholding justice” or “the law does not punish the masses”) reduces guilt and thus amplifies trolling.

However, this study did not identify a moderating role of perceived anonymity between self-esteem and online trolling. This indicates that perceived anonymity fails to additionally modify the self-esteem–online trolling association. Regardless, core of the I^3 model lies in the dynamic interplay of instigation, impellance, and inhibition (Finkel, 2014; Finkel & Hall, 2018). For a situational factor to exert a moderating effect that amplifies or attenuates a relationship, the inhibition process itself must possess a certain degree of moderating potential (Finkel,

2014; Finkel & Hall, 2018). However, as inhibition, the effect of self-esteem on online trolling may exhibit a “threshold effect.” For individuals with high self-esteem, (Campbell, 1990). Also, the disinhibiting effect induced by perceived anonymity may struggle to override the individual’s internal self-regulatory standards (Jannati et al., 2020). The behavioral motivation stemming from insufficient self-esteem may be driven by compensatory needs and identity processes for self-esteem repair (Tracy & Robins, 2003; Sest & March, 2017). However, once a group identity is activated a norm-driven deindividuation may reduce the additional influence of anonymity (Postmes & Spears, 1998).

We observed a significant gender difference in online trolling, whereby males exhibited elevated scores relative to females, corroborating prior research (Lee & Soonah., 2023). In addition, there were significant differences in online trolling by place of residence, with individuals living in towns and cities having significantly higher online trolling scores than those living in rural areas, consistent with previous research (Cabrera et al., 2024). This may be due to the fact that internet penetration and usage are higher in urban areas than in rural areas, and the popularity and use of online technologies may increase the risk of individuals becoming online trolls.

Implications for research and practice

Research on the relationship and underlying mechanisms between celebrity worship and online trolling provides actionable insights for mitigation. Targeted intervention strategies can be implemented from four aspects: First, psychologists may adopt CBT-informed “self-concept reconstruction” techniques for boosting self-esteem, delivering tailored group counseling to foster self-worth that is independent of celebrity evaluations (Kolubinski et al., 2018), alongside scenario-based role-plays to clarify communication boundaries and reduce responsibility diffusion under anonymity. Second, platform designers can optimize technical governance by piloting at-posting moral reminders for high-risk comments and adding “diverse perspective” features on celebrity topics to increase exposure to rational discourse, with ongoing evaluation (Windisch et al., 2022), given that homogeneous discussions can foster polarization (Van Swol, 2009). Third, policymakers may enforce collaborative supervision via targeted policies: requiring agencies and platforms to issue rational-fandom guidelines and strengthen moderator training with graduated enforcement, and establishing cross-platform coordination mechanisms for repeat offenders. Fourth, educational institutions and media outlets can deliver scenario-based guidance: colleges can integrate digital literacy and online civility training into curricula, while mainstream media and platforms promote “rational fandom” initiatives and misinformation-refutation content.

Limitations and future directions

This study had the following limitations: First, although it provided empirical evidence of the association between celebrity worship and online trolling, it did not differentiate between groups with distinct worship types.

Future research could compare trolling tendencies across clearly defined celebrity/fandom types (e.g., “traffic celebrities” vs. “talented idols”) and test whether key associations differ across types. Second, as a cross-sectional design, this study could not infer causality or capture temporal changes in celebrity worship and trolling behaviors. Future research could use multi-wave longitudinal designs to examine temporal ordering and cross-time effects among key variables, and complement this with diary/experience-sampling methods to capture short-term fluctuations, thereby strengthening dynamic inference and causal interpretation. Third, the sample surveyed in this study is predominantly drawn from urban populations. Future studies could use stratified sampling across urban–rural regions and socioeconomic contexts and test whether key effects are consistent across regions, thereby clarifying whether the proposed mechanisms are general or context-dependent. Fourth, Future studies could implement standardized mode differentiation (e.g., fully paper-based offline responses vs. web-based online responses) and test mode equivalence to evaluate whether findings are robust across collection modes and to strengthen external validity. Finally, the extent of social desirability is unknown, as participants might have concealed their true attitudes to present themselves more positively. Future research could mitigate this bias by triangulating self-reports with behavioral indicators (e.g., platform activity logs or simulated commenting tasks) and alternative reporting approaches.

Conclusion

Drawing upon the I^3 framework (Finkel, 2014), the current study tested a model linking celebrity worship to online trolling among college students, with self-esteem and moral disengagement as sequential mediators and perceived anonymity as a moderator. Findings pointed to a significant positive correlation between the two focal variables. Both self-esteem and moral disengagement independently mediated this relationship, and their sequential chain mediation further explained the mechanism. Furthermore, perceived anonymity moderated the direct and indirect pathways between celebrity worship and online trolling: among high-perceived-anonymity individuals, the positive predictive effects of celebrity worship and moral disengagement on trolling were significantly amplified compared to low-anonymity counterparts. As a critical trigger of online trolling, celebrity worship warrants deeper exploration. Investigating celebrity worship and trolling behaviour underlying psychological mechanisms offers novel perspectives for mitigating trolling behaviors and suggest innovative intervention strategies to enhance healthy cyber-governance frameworks.

Acknowledgement: Funding support for this study is gratefully acknowledged, and we are indebted to all participants for their engagement.

Funding Statement: This work received funding from the National Social Science Foundation of China (grant number: 23BSH143).

Author Contributions: The authors have confirmed their contributions to the paper as follows: study design: Fang Li, Mengni Yu and Wenyu Zeng; data collection: Mengni Yu and Wenyu Zeng; analysis and interpretation of results: Mengni Yu, Wenyu Zeng and Yihong Yang; drafting the manuscript: Mengni Yu, Wenyu Zeng, Yihong Yang, Fang Li and Biyun Wu. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data supporting this research are available from the corresponding author upon request.

Ethics Approval: This study has been approved by the Ethics Committee of Sichuan Normal University (IRB No. SCNU-230904). All participants furnished voluntary written informed consent following thorough briefing on the study procedures and their entitlements, including the right to withdraw at any time.

Conflicts of Interest: The authors declare no conflicts of interest.

Supplementary Materials: The supplementary material is available online at <https://www.techscience.com/doi/10.32604/jpa.2026.068011/s1>.

References

- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193–209. https://doi.org/10.1207/s15327957pspr0303_3
- Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, 71(2), 364. <https://doi.org/10.1037/0022-3514.71.2.364>
- Cabrera, M. C., Larrañaga, E., & Yubero, S. (2024). Bullying/cyberbullying in secondary education: A comparison between secondary schools in rural and urban contexts. *Child and Adolescent Social Work Journal*, 41(4), 617–631. <https://doi.org/10.1007/s10560-022-00882-0>
- Campbell, J. D. (1990). Self-esteem and clarity of the self-concept. *Journal of Personality and Social Psychology*, 59(3), 538. <https://doi.org/10.1037/0022-3514.59.3.538>
- Chu, X. W., Zhou, Z. K., & Fan, C. Y. (2020). Relationship between peer rejection and cyberbullying: The mediating role of anger and the moderating role of perceived anonymity. *Psychological Development and Education*, 36(5), 584–593. (In Chinese). <https://doi.org/10.16187/j.cnki.issn1001-4918.2020.05.10>
- Cui, J., Hong, J. Z., Fang, Y., Chu, X. W., & Wu, K. (2022). The effect of celebrity worship on problematic internet use in college students: Chain mediating roles of the desire for fame and fear of missing out. *Chinese Journal of Clinical Psychology*, 30(6), 1287–1291. (In Chinese). <https://doi.org/10.16128/j.cnki.1005-3611.2022.06.005>
- Finkel, E. J. (2014). The I^3 model: Metatheory, theory, and evidence. *Advances in Experimental Social Psychology*, 49, 1–104. <https://doi.org/10.1016/B978-0-12-800052-6.00001-9>
- Finkel, E. J., & Hall, A. N. (2018). The I^3 model: A metatheoretical framework for understanding aggression. *Current*

- Opinion in Psychology*, 19, 125–130. <https://doi.org/10.1016/j.copsyc.2017.03.013>
- Fox, J., & Vendemia, M. A. (2016). Selective self-presentation and social comparison through photographs on social networking sites. *Cyberpsychology, Behavior, and Social Networking*, 19(10), 593–600. <https://doi.org/10.1089/cyber.2016.0248>
- Gasser, L., & Keller, M. (2009). Are the competent the morally good? Perspective taking and moral motivation of children involved in bullying. *Social Development*, 18(4), 798–816. <https://doi.org/10.1111/j.1467-9507.2008.00516.x>
- Giles, D. C., & Maltby, J. (2004). The role of media figures in adolescent development: Relations between autonomy, attachment, and interest in celebrities. *Personality and Individual Differences*, 36(4), 813–822. [https://doi.org/10.1016/S0191-8869\(03\)00154-5](https://doi.org/10.1016/S0191-8869(03)00154-5)
- Hauben, M., & Hauben, R. (1997). *Netizens: On the History and Impact of Usenet and the Internet*. Los Alamitos, CA: IEEE Computer Society Press.
- He, F. C., & Xu, Y. (2023). Relationship between celebrity worship and cyberbullying in middle school students: The chain mediating role of self-esteem and defense styles. *China Journal of Health Psychology*, 31(5), 790–796. (In Chinese). <https://doi.org/10.13342/j.cnki.cjhp.2023.05.028>
- Jannati, Y., Nia, H. S., Froelicher, E. S., Goudarzian, A. H., & Yaghoobzadeh, A. (2020). Self-blame attributions of patients: A systematic review study. *Central Asian Journal of Global Health*, 9(1), e419. <https://doi.org/10.5195/cajgh.2020.419>
- Jung, Y., Song, H., & Vorderer, P. (2012). Why do people post and read personal messages in public? The motivation of using personal blogs and its effects on users' loneliness, belonging, and well-being. *Computers in Human Behavior*, 28(5), 1626–1633. <https://doi.org/10.1016/j.chb.2012.04.001>
- Kolubinski, D. C., Frings, D., Nikčević, A. V., Lawrence, J. A., & Spada, M. M. (2018). A systematic review and meta-analysis of CBT interventions based on the Fennell model of low self-esteem. *Psychiatry Research*, 267, 296–305. <https://doi.org/10.1016/j.psychres.2018.06.025>
- Lee, G., & Soonah, A. (2023). Anonymity and gender effects on online trolling and cybervictimization. *Journal of Cybersecurity Education, Research and Practice*, 2023(1), 5. <https://doi.org/10.32727/8.2023.14>
- Li, J. L., Shan, F. P., Huang, J. N., & Huang, L. Y. (2023). Relationship between idolatry and perceived social support in secondary school students: The mediating role of peer attachment and life satisfaction. *Sichuan Mental Health*, 36(4), 359–365. (In Chinese). <https://doi.org/10.11886/scjsws20230320003>
- Li, F., Tang, X., Ge, X., Yu, M. N., & Wang, S. Q., et al. (2024). Is mindfulness associated with lower online trolling among adolescents? Mediating effects of self-esteem and depression and moderating effect of dark personality traits. *Journal of Psychology in Africa*, 34(3), 293–298. <https://doi.org/10.1080/14330237.2024.2363698>
- Maltby, J., Day, L., McCutcheon, L. E., Martin, M. M., & Cayanus, J. L. (2004). Celebrity worship, cognitive flexibility, and social complexity. *Personality and Individual Differences*, 37(7), 1475–1482. <https://doi.org/10.1016/j.paid.2004.02.004>
- McCutcheon, L. E., Lange, R., & Houran, J. (2002). Conceptualization and measurement of celebrity worship. *British Journal of Psychology*, 93(1), 67–87. <https://doi.org/10.1348/000712602162454>
- Nitschinsk, L., Tobin, S. J., & Vanman, E. J. (2022). The disinhibiting effects of anonymity increase online trolling. *Cyberpsychology, Behavior, and Social Networking*, 25(6), 377–383. <https://doi.org/10.1089/cyber.2022.0005>
- Ouvrein, G., De Backer, C. J., & Vandebosch, H. (2018). Online celebrity aggression: A combination of low empathy and high moral disengagement? The relationship between empathy and moral disengagement and adolescents' online celebrity aggression. *Computers in Human Behavior*, 89, 61–69. <https://doi.org/10.1016/j.chb.2018.07.029>
- Peng, W. B., Qiu, X. T., Liu, D. Z., & Wang, P. (2010). Revision of the scale of celebrity worship. *Psychological Development and Education*, 26(5), 543–548. (In Chinese). <https://doi.org/10.16187/j.cnki.issn1001-4918.2010.05.002>
- Pirzade, M., Peyvastegar, M., & Griffiths, M. D. (2024). Celebrity worship among adolescents is driven by neuroticism, avoidant identity style, and need to belong. *The Journal of Genetic Psychology*, 185(6), 463–476. <https://doi.org/10.1080/00221325.2024.2331731>
- Postmes, T., & Spears, R. (1998). Deindividuation and antinormative behavior: A meta-analysis. *Psychological Bulletin*, 123(3), 238. <https://doi.org/10.1037/0033-2909.123.3.238>
- Putri, S. D. G., Purnomo, E. P., & Khairunissa, T. (2024). Echo chambers and algorithmic bias: The homogenization of online culture in a smart society. *SHS Web of Conferences*, 202, 05001. <https://doi.org/10.1051/shsconf/202420205001>
- Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.
- Runions, K. C., & Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychology, Behavior, and Social Networking*, 18(7), 400–405. <https://doi.org/10.1089/cyber.2014.0670>
- Saulnier, L., & Krettenauer, T. (2023). Internet impropriety: Moral identity, moral disengagement, and antisocial online behavior within an early adolescent to young adult sample. *Journal of Adolescence*, 95(2), 264–283. <https://doi.org/10.1002/jad.12112>
- Sest, N., & March, E. (2017). Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences*, 119, 69–72. <https://doi.org/10.1016/j.paid.2017.06.038>
- Singh, S. (2024). The role of anonymity and online disinhibition in enabling trolling behavior. *Sudarshan Research Journal*, 2(3), 73.
- Stepanyan, A. (2019). *New media literacy: Celebrity worship, Instagram, and self-esteem* (Doctoral dissertation). California State University.
- Sude, D. J., & Dvir-Gvirsman, S. (2023). Different platforms, different uses: Testing the effect of platforms and individual differences on perception of incivility and self-reported uncivil behavior. *Journal of Computer-Mediated Communication*, 28(2), zmac035. <https://doi.org/10.1093/jcmc/zmac035>
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, 7(3), 321–326. <https://doi.org/10.1089/1094931041291295>
- Tajfel, H., Turner, J. C., Austin, W. G., & Worchel, S. (1979). An integrative theory of intergroup conflict. *Organizational Identity: A Reader*, 56(65), 9780203505984–16. <https://doi.org/10.1093/oso/9780199269464.003.0005>
- Tesser, A., Millar, M., & Moore, J. (1988). Some affective consequences of social comparison and reflection processes: The pain and pleasure of being close. *Journal of Personality and Social Psychology*, 54(1), 49. <https://doi.org/10.1037/0022-3514.54.1.49>

- Thomaes, S., Bushman, B. J., Stegge, H., & Olthof, T. (2008). Trumping shame by blasts of noise: Narcissism, self-esteem, shame, and aggression in young adolescents. *Child Development, 79*(6), 1792–1801. <https://doi.org/10.1111/j.1467-8624.2008.01226.x>
- Tracy, J. L., & Robins, R. W. (2003). “Death of a (narcissistic) salesman:” An integrative model of fragile self-esteem. *Psychological Inquiry, 14*(1), 57–62.
- Van Swol, L. M. (2009). Extreme members and group polarization. *Social Influence, 4*(3), 185–199. <https://doi.org/10.1080/15534510802584368>
- Wachs, S., & Wright, M. F. (2018). Associations between bystanders and perpetrators of online hate: The moderating role of toxic online disinhibition. *International Journal of Environmental Research and Public Health, 15*(9), 2030. <https://doi.org/10.3390/ijerph15092030>
- Windisch, S., Wiedlitzka, S., Olaghere, A., & Jenaway, E. (2022). Online interventions for reducing hate speech and cyberhate: A systematic review. *Campbell Systematic Reviews, 18*(2), e1243. <https://doi.org/10.1002/cl2.1243>
- Wright, M. F. (2013). The relationship between young adults’ beliefs about anonymity and subsequent cyber aggression. *Cyberpsychology, Behavior, and Social Networking, 16*(12), 858–862. <https://doi.org/10.1089/cyber.2013.0009>
- Yang, J. P., & Wang, X. C. (2011). Interparental conflict and aggressive behavior: The mediating effect of the moral disengagement. *Psychological Development and Education, 27*(5), 498–505. (In Chinese). <https://doi.org/10.16187/j.cnki.issn1001-4918.2011.05.001>
- Zezulka, L. A., & Seigfried-Spellar, K. C. (2016). Differentiating cyberbullies and internet trolls by personality characteristics. *The Journal of Digital Forensics, Security and Law, 11*(3), 7–26. <https://doi.org/10.58940/1558-7223.1410>
- Zhang, L., Li, Y., & Song, X. (2025). Interpreting cyberbullying within fan culture: The relationship between celebrity worship and cyberbullying. *Social Behavior and Personality: An International Journal, 53*(2), e13895. <https://doi.org/10.2224/sbp.13895>