

How Social Media Algorithms May Exacerbate the Identifiable Victim Effect and Empathy Fatigue among Adolescents: A Critical Narrative Review

Zhengyou Zhou

*Faculty of Science, The University of Sydney, Sydney, Australia
zzho0757@uni.sydney.edu.au*

Abstract. Algorithmic curation repeatedly exposes teenagers to vivid narratives of individual pain. This review examines whether the participation-based recommendation system will strengthen the identifiable victim effect (IVE), whether identifiable individuals produce stronger emotional or helping responses than statistical victims, and whether repeated exposure will lead to empathy fatigue. The evidence integrates moral psychology, recommendation system audit, social learning research, developmental neuroscience, adolescent media research and alternative trauma research. The literature supports two interrelated propositions, but does not support a deterministic causal chain. First of all, the ranking and feedback cycle based on participation give priority to the emotional and moral prominent content, so that the identifiable victim narrative may be amplified structurally. Second, repeated exposure may lead to habituation characterized by blunted emotion, or overload characterized by distress and avoidance; existing studies rarely measure affect, helping behavior, and exposure longitudinally enough to distinguish these pathways. Adolescence may increase susceptibility because social emotional responses, peer sensitivity, and regulatory capacity are still developing, although direct evidence for adolescents is lacking. This review further points out that self-efficacy is a possible moderator: amplified feeling without a reliable avenue for action may turn worry into helplessness. In general, algorithmic amplification is theoretically reasonable and indirectly supported, but the claim that it leads to adolescent empathy fatigue has not been confirmed. Research should directly link feed level exposure, victim identifiability, impact, efficacy, and help over time.

Keywords: adolescent social media, identifiable victim effect, recommender systems, empathy fatigue, vicarious traumatization

1. Introduction

People usually respond more strongly to a specific victim than to the pain described statistically. Even minimal certainty about who will receive assistance can increase helping [1]. For an identified victim, the effect is also significantly stronger than for an identified group [2]. This asymmetry, often referred to as the identifiable victim effect (IVE), directs attention and concern to vivid

individuals, even when larger, less obvious groups face greater needs. A recent comprehensive study has linked this effect to emotional saliency, psychological distance and perceived impact [3].

The media environment in which teenagers suffer is fundamentally different from the limited environment used in classical IVE experiments. Personalized feeds provide a continuous sequence of stories, images, and videos, selected in part by predictive engagement. A pre-registered audit found that emotional and disruptive materials were amplified in the engagement-based timeline relative to the reverse-chronological baseline drawn from the same followed accounts [4]. Social learning research further shows that algorithmic rewards not only shape what users see, but also shape what users and organizations learn [5]. Therefore, identifiability may shift from discrete experimental operations to repetitive features of information environment.

Repeated exposure has no definite psychological consequences. The audit of the youth-oriented recommendation system recorded escalating exposure to, or normalization of, harmful content, while surveys and psychometric studies linked the use of the recommendation system or frequent traumatic media exposure to vicarious trauma and anxiety [6-9]. Qualitative evidence obtained from girls exposed to war related videos describes intense pain and deliberate avoidance [10]. These findings make the empathy fatigue related to the algorithm credible, but there is no complete test of the sequence from victim identifiability, through amplification by the algorithm, to changes in teenagers' emotions and helping behavior.

Adolescence may be an important environment, because sociocultural sensitivity, peer evaluation and self-regulation are constantly changing during this period, and the brain system related to social cognition is also developing [11, 12]. However, development sensitivity should not be considered as evidence of exceptional vulnerability. There is a lack of direct evidence linking adolescent IVE with recommender-system exposure.

Therefore, this review asks whether participation-based recommendation systems can amplify the narratives of identifiable victims, and whether repeated contact can contribute to adolescent empathy fatigue. Three related but separable propositions are proposed: a selection proposition, in which engagement ranking favors characteristics related to identifiability; a saturation proposition, in which repeated exposure will lead to habituation or emotional overload; and a developmental proposition, in which adolescents' social emotional sensitivity and regulatory conditions are still developing. This review uses a critical narrative synthesis of moral psychology, recommendation system audit, social learning, developmental neuroscience, adolescent media research and alternative trauma research. This review is not a systematic review, nor does it estimate a pooled effect. It assesses the coherence, evidence support, and limitations of causal arguments assembled throughout the research tradition.

2. Conceptual foundations

2.1. The identifiable victim effect

IVE is a difference in emotional or prosocial responses that results from presenting specific victims as identifiable rather than statistically describing equivalent needs. Classical research has established two core features: identifiability can play a role without a large number of personal details, and the effect is greatest for one victim rather than a group [1, 2]. Recognizability overlaps with vividness and emotionality, but it is not synonymous with either. A dramatic event can be emotional without identifying the victim, while a designated recipient can be identifiable without graphic display.

Three mechanisms are most relevant to this review: vivid individuals reduce psychological distance and recruit rapid emotional processing; helping a person can feel more complete and

effective than making an insignificant contribution to a big problem; the single victim provides a coherent attention goal [3]. These mechanisms identify observable content features - faces or names, coherent personal narratives, and significant opportunities for help - which can be indirectly rewarded by the ranking system. They themselves do not prove that the platform detects identifiability, or that amplified emotions become prosocial behaviors. This difference constitutes the latter analysis.

2.2. Distinguishing processes grouped as empathy fatigue

Empathy fatigue is used here only as an umbrella term. At least three processes differ in time scale, internal state, and behavioral prediction, as summarized in Table 1.

Table 1. Processes commonly grouped as empathy fatigue

Process	Temporal pattern	Core mechanism	Expected affect and behavior
Compassion fade or collapse	Within one episode	Scope insensitivity or motivated down-regulation	Weaker concern or helping for many victims than for one
Desensitization or habituation	Across repeated episodes	Threshold shift and normalization	Blunted affect with indifference or reduced helping
Overload or vicarious traumatization	Cumulative exposure	Emotional flooding and regulatory failure	High distress with avoidance or withdrawal

Compassion fade involves a response within a single event. An influential explanation regards the reduction of sympathy of many victims as motivated emotion regulation, but a high-intensity replication does not reproduce the key interactions and identified measurement problems in the original paradigm [13, 14]. Recent evidence also shows that recognizability can change emotions more reliably than donation behavior, while meta-analysis estimates that the overall correlation between recognizability and help is very small [15, 16]. Therefore, empathy regression should not be regarded as a fixed mechanism of long-term fatigue.

Desensitization and overload are associated with repeated exposure. The former indicates that the emotional response gradually weakens; the latter indicates sustained or aggravated pain, leading to avoidance. Audits provide evidence on the supply and sequencing of harmful content [6, 7]. While trauma-related research provides evidence on pain and withdrawal [8-10]. Because the reduction of participation or help may be accompanied by numbness or overwhelming emotion, behavioral withdrawal itself cannot determine its operating mechanism.

3. Engagement-based ranking and identifiability

3.1. Emotional amplification as a selection pressure

The clearest evidence of algorithm selection comes from the comparison between the personalized timeline based on participation and the reverse timeline constructed by the same follow-up account. Participation ranking magnified anger, sadness, anxiety, happiness, ideological tendency and hostility outside the group; Exposure also worsened attitudes towards political expatriate groups. Participants do not like the political content chosen for them [4]. Keeping the followed accounts

unchanged strengthens the inference that ranking (not just users' previous account selections) changes exposure.

Relevant inferences are limited but important. Participation goals do not require a clear category of victims to reward the common features of identifiable victim narratives: visible people, emotional clarity, moral interests, and compact stories. Content with these features may generate signals for clicking, viewing time, and using the shared ranking system. Therefore, identifiability may become an indirect selection pressure. Audit measures emotional and divisive rather than the identifiability of victims, so this is still a theoretically based inference rather than a proven platform effect.

3.2. Reciprocal feedback between ranking and production

Selection works together with adaptation. The PRIME framework proposes that human social learning is biased toward prestigious, in-group, moral, and emotional information, while engagement-optimizing algorithms use the same biases [5]. Experimental and observational evidence further shows that social feedback in the form of liking and sharing will increase the subsequent expression of moral anger [17]. A broader review also concluded that the moral and emotional content is constantly amplified, although the downstream social impact is still difficult to determine [18].

Applied to the victim's narration, this feedback loop predicts a simple side effect: charities, news organizations and individual users know which frames travel, and then copy these frameworks. A single victim provides a clearer unit of attention than the overall statistics, so repeated feedback may make the pain more and more obvious through personalized stories. No study cited here directly tests this adaptation of victim content; this statement identifies a possible mechanism to measure rather than an established outcome.

3.3. Limits of the selection evidence

The existing evidence cannot prove that the input ranking enlarges the IVE of teenagers. The most powerful audit conducted short-term research on adults on a platform, mainly relying on self-reported attitudes, without measuring helping behavior [4]. Social learning evidence involves moral emotional expression, not identifiable victims, and is not specifically evaluated for minors [5, 17]. Emotionality, moral saliency, vividness, and identifiability overlap, but treating them as interchangeable exaggerates the evidence. Direct testing must encode the identifiability of victims at the project level, and link exposure with emotion and corresponding prosocial results.

4. From amplification to fatigue: rival pathways

4.1. Habituation and normalization

According to the habituation theory, repeated exposure increases the threshold of significance, making previously shocking materials appear ordinary. An audit report combining interviews, school leaders' discussions and analysis of more than 1000 videos said that the four vulnerable profiles received increasing "doses" of harmful ideological material within seven days, and harmful information was normalized through entertainment [6]. Another survey using more than 30 automated accounts simulating 13-year-old children found that feeds quickly focused on mental health and potentially harmful videos after receiving signals of interest [7].

These studies show that the recommender system has the ability to amplify and redeliver content; they did not show a decline in empathy. The first is about misogyny and ideology, rather than the victim's narration, which partly relies on simulated personal data. The second used simulated and manually operated accounts in an advocacy audit. Both can determine what the system can provide, but neither can track teenagers' emotional response or helping behavior. Therefore, the evidence of exposure escalation is stronger than the evidence of habituation.

4.2. Emotional overload and vicarious traumatization

The competing pathway is emotional overload rather than numbness. In a survey of 1,317 Chinese college students following the MU5735 crash, social media use predicted vicarious trauma ($\beta=0.180$), reliance on recommendation systems independently predicted trauma ($\beta=0.172$), and peer communication was the strongest predictor ($\beta=0.303$); dependence on recommendation systems also partially mediated the association between social media use and mental trauma [8]. The cross-sectional design and a sample of young adults prevented causal or adolescent-specific conclusions.

Psychometric research on young people found that media substitution trauma and war-related concerns can predict general and future oriented anxiety, and emphasized that exposure frequency is particularly relevant [9]. The most age-related evidence comes from interviews with 15 girls aged 15-18 who saw videos related to the war on the Internet. Participants described severe physical distress, unnecessary exposure, avoidance, and dependence on peers; Content warnings cannot reliably block viewing [10]. These findings are more closely related to overload than habituation, but the small qualitative sample and lack of feed logs or measures of helping limit generalization.

4.3. Why the pathways remain empirically indistinguishable

The laboratory evidence does not resolve the conflict. Motivated downward adjustment suggests that people reduce compassion when anticipating expensive needs [13]. However, in a pre-registered high-powered replication with 1177 participants, the central experimental results were not reproduced [14]. In addition, the identifiability of victims can increase sympathy without increasing donations [15]. Therefore, feed may produce repeated emotional arousal without a corresponding sense of effective action.

Habituation and overload can produce the same observable behavior - disengagement from painful content - despite opposite emotional states. Supply reviews infer responses without measurement, while trauma studies measure pain without project level exposure or prosocial action. In the evidence reviewed, no study tracked the identifiability, exposure order, emotion, perceived efficacy and helping behavior in the same adolescent participants. Therefore, "adolescent empathy fatigue" should be regarded as an unresolved hypothesis rather than a documented effect of the recommendation system.

5. Adolescence as a conditional vulnerability

5.1. Developmental rationale

The debate on development is based on convergence, not on the general idea of teenagers' impulsivity. The social cognitive system involving the medial prefrontal cortex and the temporal parietal junction continues to develop throughout adolescence, alongside mentalizing, peer

sensitivity and emotion regulation [12]. The development model also describes a period in which social emotional responses may exceed regulatory control [11]. Therefore, while regulation and self-definition are still developing, vivid victim narratives and visible peer reactions may be particularly prominent.

A large body of longitudinal evidence warns against regarding adolescence as a single unified window. The analysis of 17,409 participants aged 10-21 years identified age- and gender-specific periods related to social media use and life satisfaction, but the effects were small and bidirectional, and exposure was measured using self-reported time [19]. Online time cannot reveal whether two teenagers have experienced the same amount, order or identifiable pain. This study supports developmental heterogeneity rather than causality from algorithmic exposure to victim content.

Peer cues can also mitigate rather than merely increase risk. The experiment on fund-raising information found that cues suggestive of others' support improved the attitude and willingness to help multiple victims through perceived impact and self-efficacy [20]. Such clues can expand attention beyond a victim and make action seem consequential. Because this study did not test adolescents or longitudinal feed exposure, the increased sensitivity of adolescents to these cues is still a reasonable and testable proposition.

5.2. Limits of the developmental claim

Three limitations preclude strong development conclusions. First of all, few studies have directly tested whether puberty is a sensitive period of lasting effects, while developmental neuroscience samples are mainly concentrated in the western population [11]. Second, the classic IVE literature mainly relies on adult samples and assumptions or laboratory decisions, while adolescent social media research usually measures time use and well-being rather than victim recognition. Third, simulated accounts are usually used in auditing, and the most direct evidence of teenagers' actual contact is a small qualitative study [10]. Therefore, the proposition that algorithms amplify IVE among adolescents is conditional and indirect at present.

6. An integrative feedback account

6.1. Coupling exposure, action, and efficacy

The choice and saturation propositions are better understood as a feedback system rather than two additive effects. Participation ranking is conducive to emotional signals, producers adapt to algorithms and social rewards, and teenagers may repeatedly encounter morally salient cases [4, 5, 17]. Exposure then generates behavior data - viewing, sharing, skipping, or exiting - that inform subsequent rankings. The system can adjust its supply without knowing whether the withdrawal reflects apathy, distress or satisfaction.

The central moderator is perceived efficacy. Recognizability can produce emotions more consistently than help, while vicarious-trauma research places helplessness near the center of overload [9, 15]. Therefore, when teenagers are unable to turn their concerns into corresponding actions, repeated arousal may be costly. On the contrary, strengthening social cues of perceived impact and self-efficacy can increase support for multiple victims [20]. The most reasonable synthesis is not that exposure to pain will inevitably exhaust empathy, but that amplified feeling without a reliable avenue for action may encourage distress, avoidance or disengagement.

6.2. Objections and alternative explanations

Three objections limit this statement. First of all, on average, behavioral IVE is small and usually does not exist, so algorithmic amplification may change emotions more than resource allocation [15, 16]. Second, platform cues can expand attention and restore efficiency, which shows that single-victim bias is not an inevitable result of digital design [20]. Third, the causal direction has not yet been resolved: associations with social media are bidirectional, and teenagers may actively seek graphic content through curiosity, or contact graphic content through ranking [10, 19]. These alternatives do not eliminate the impact of the algorithm, but they exclude deterministic declarations.

7. Ethical and social implications

The first ethical issue is distribution. The information environment optimized for individuals with clear emotions may allocate attention according to visibility rather than severity or scale [3]. Slow, diffuse and statistically representative hazards - including chronic poverty, displacement and preventable diseases - are not compatible with narrative features that generate rapid participation. Therefore, algorithmic selection can replicate sympathetic inequality even if there is no clear intention to privilege a particular victim.

The second problem is teenagers' autonomy. The participation-based system can use user behavior to predict the content users will engage with, but such behavior does not necessarily indicate preferred content [4]. When the material describes human suffering and the user is a minor, behavioral prediction is a weak proxy for considered preference. Inevitable graphic exposure and ineffective warnings further indicate that adolescents may suffer when they lack meaningful control and adequate support [10].

The third issue is accountability. Some of the most informative evidence relies on advocacy audits or simulated accounts, because independent researchers cannot obtain real exposure data [6, 7]. The opacity of the platform limits the ability to distinguish between algorithm selection and user selection, as well as the ability to evaluate the age-specific outcomes, which is an issue highlighted in the broader comments on social media and ethics [18]. Therefore, independent and privacy protected access to project level exposure and ranking data is necessary for credible assessment and appropriate protection.

8. Conclusion

The idea that social media algorithms aggravate the identifiable victim effect and lead to empathy fatigue in adolescents is consistent in theory, but has not been confirmed. The ranking based on participation tends to favor emotional and personalized content, while social feedback encourages producers to copy the format of communication. Repeated exposure may lead to habituation or emotional overload, but the current evidence cannot distinguish these pathways, nor can it show that either changes adolescents' helping behavior.

Adolescence is best seen as a conditional environment, rather than a period of exceptional vulnerability to algorithms. The development of social sensitivity, peer responsiveness and regulatory ability may affect responses to victim narratives, but the necessary direct evidence is lacking. The most powerful comprehensive insight is that perceived efficacy may determine whether amplified attention supports action or instead produces helplessness and withdrawal. A decisive test must track actual feed exposure over time and measure victim identifiability, affect, perceived

efficacy, and corresponding helping behavior in the same adolescents. Before that, causal language should be restrained, and ethical concerns should focus on transparent exposure, meaningful user control and unequal distribution of visibility.

References

- [1] Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty*, 26(1), 5-16. <https://doi.org/10.1023/A:1022299422219>
- [2] Kogut, T., & Ritov, I. (2005). The identified victim effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, 18(3), 157-167. <https://doi.org/10.1002/bdm.492>
- [3] Zhou, Z. (2026). Psychological mechanisms and impacts of the identifiable victim effect. *SHS Web of Conferences*, 235, Article 03012. <https://doi.org/10.1051/shsconf/202623503012>
- [4] Milli, S., Carroll, M., Wang, Y., Pandey, S., Zhao, S., & Dragan, A. D. (2025). Engagement, user satisfaction, and the amplification of divisive content on social media. *PNAS Nexus*, 4(3), Article pgaf062. <https://doi.org/10.1093/pnasnexus/pgaf062>
- [5] Brady, W. J., Jackson, J. C., Lindström, B., & Crockett, M. J. (2023). Algorithm-mediated social learning in online social networks. *Trends in Cognitive Sciences*, 27(10), 947-960. <https://doi.org/10.1016/j.tics.2023.06.008>
- [6] Regehr, K., Shaughnessy, C., Zhao, M., Cambazoglu, I., Turner, A., & Shaughnessy, N. (2025). Normalizing toxicity: The role of recommender algorithms for young people's mental health and social wellbeing. *Frontiers in Psychology*, 16, Article 1523649. <https://doi.org/10.3389/fpsyg.2025.1523649>
- [7] Amnesty International. (2023, November 7). Driven into darkness: How TikTok's 'For You' feed encourages self-harm and suicidal ideation. <https://www.amnesty.org/en/documents/POL40/7350/2023/en/>
- [8] Li, K., Li, J., & Li, Y. (2024). The effects of social media usage on vicarious traumatization and the mediation role of recommendation systems usage and peer communication in China after the aircraft flight accident. *European Journal of Psychotraumatology*, 15(1), Article 2337509. <https://doi.org/10.1080/20008066.2024.2337509>
- [9] Regnoli, G. M., Tiano, G., & De Rosa, B. (2025). Media exposure and vicarious trauma: Italian adaptation and validation of the Media Vicarious Traumatization Scale and its impact on young adults' mental health in relation to contemporary armed conflicts. *European Journal of Investigation in Health, Psychology and Education*, 15(9), Article 184. <https://doi.org/10.3390/ejihpe15090184>
- [10] Dolev-Cohen, M., Kakun, P., & Toizner, B. (2026). Digital wounds: The coping experiences of adolescent girls exposed to snuff videos online. *Child and Adolescent Social Work Journal*, 43(4), 1303-1315. <https://doi.org/10.1007/s10560-026-01087-5>
- [11] Cheng, T. W., Mills, K. L., & Pfeifer, J. H. (2024). Revisiting adolescence as a sensitive period for sociocultural processing. *Neuroscience & Biobehavioral Reviews*, 164, Article 105820. <https://doi.org/10.1016/j.neubiorev.2024.105820>
- [12] Andrews, J. L., Ahmed, S. P., & Blakemore, S.-J. (2021). Navigating the social environment in adolescence: The role of social brain development. *Biological Psychiatry*, 89(2), 109-118. <https://doi.org/10.1016/j.biopsych.2020.09.012>
- [13] Cameron, C. D., & Payne, B. K. (2011). Escaping affect: How motivated emotion regulation creates insensitivity to mass suffering. *Journal of Personality and Social Psychology*, 100(1), 1-15. <https://doi.org/10.1037/a0021643>
- [14] Hagman, W., Tinghög, G., Dickert, S., Slovic, P., & Västfjäll, D. (2022). Motivated down-regulation of emotion and compassion collapse revisited. *Frontiers in Psychology*, 13, Article 801150. <https://doi.org/10.3389/fpsyg.2022.801150>
- [15] Moche, H., Karlsson, H., & Västfjäll, D. (2024). Victim identifiability, number of victims, and unit asking in charitable giving. *PLOS ONE*, 19(3), Article e0300863. <https://doi.org/10.1371/journal.pone.0300863>
- [16] Lee, S., & Feeley, T. H. (2016). The identifiable victim effect: A meta-analytic review. *Social Influence*, 11(3), 199-215. <https://doi.org/10.1080/15534510.2016.1216891>
- [17] Brady, W. J., McLoughlin, K., Doan, T. N., & Crockett, M. J. (2021). How social learning amplifies moral outrage expression in online social networks. *Science Advances*, 7(33), Article eabe5641. <https://doi.org/10.1126/sciadv.abe5641>
- [18] Van Bavel, J. J., Robertson, C. E., del Rosario, K., Rasmussen, J., & Rathje, S. (2024). Social media and morality. *Annual Review of Psychology*, 75, 311-340. <https://doi.org/10.1146/annurev-psych-022123-110258>
- [19] Orben, A., Przybylski, A. K., Blakemore, S.-J., & Kievit, R. A. (2022). Windows of developmental sensitivity to social media. *Nature Communications*, 13(1), Article 1649. <https://doi.org/10.1038/s41467-022-29296-3>

- [20] Chun, J. W., & Lee, A. R. (2026). Going beyond the single identified victim effect: Exploring social media cues to promote supportive intentions for multiple victims in fundraising campaigns. *Humanities and Social Sciences Communications*. Advance online publication. <https://doi.org/10.1057/s41599-026-07936-z>