3rd International Conference on Education, Environment, Arts and Social Science (EEAS 2026)

Article

A Study on the Influence Mechanism of Revenge Motivation on Antisocial Behavior in Digital Social Contexts: The Mediating Role of Aggression and the Moderating Effect of the Dark Triad Personality Traits

Jia Wu[1,*]

[1] James Cook University, Singapore, Singapore
* Correspondence: Jia Wu, James Cook University, Singapore, Singapore

Abstract: The rapid development of digital social environments, including social media platforms and online forums, has introduced novel forms of interpersonal tension and antisocial behavior. Grounded in the General Aggression Model and the social cognitive theory of moral disengagement, this study proposes a moderated mediation model to examine how revenge motivation translates into antisocial behavior in digital settings. Specifically, aggression is hypothesized as a mediator in the relationship between revenge motivation and antisocial behavior, while the Dark Triad personality traits (Machiavellianism, narcissism, and psychopathy) are proposed as a moderator that amplifies this indirect effect. Using secondary data from publicly accessible and established large-scale databases, including the National Longitudinal Study of Adolescent to Adult Health and the Pew Internet Research datasets, this research employs structural equation modeling to test the hypothesized pathways. The findings are expected to reveal that revenge motivation significantly predicts antisocial behavior through increased aggression, and this mediating process is stronger among individuals with elevated Dark Triad traits. This study contributes to the theoretical understanding of cyber aggression and antisocial conduct in digital ecosystems. It also offers practical implications for educational strategies and digital citizenship programs targeting youth populations vulnerable to online victimization and retaliation cycles. By integrating personality psychology with digital behavior research, this work provides a comprehensive framework for understanding the psychological mechanisms underlying online antisocial conduct and informs evidence-based approaches to promoting healthier digital interactions.

Keywords: revenge motivation; antisocial behavior; aggression; dark triad; digital social context

Received: 20 May 2026
Revised: 24 June 2026
Accepted: 06 July 2026
Published: 11 July 2026

1. Introduction

The rapid development of digital social environments—including social media platforms, online forums, and instant messaging applications—has significantly reshaped patterns of interpersonal communication, information exchange, and relational dynamics among users across age groups. These platforms facilitate continuous connectivity and broaden opportunities for self-expression, identity exploration, and community building. However, the structural features inherent to digital interaction—such as diminished nonverbal feedback, greater psychological distance between participants, and temporal flexibility in message delivery—can influence how individuals interpret social signals and regulate emotional responses. In particular, reduced immediacy and contextual grounding may lower thresholds for impulsive reactions, thereby increasing the likelihood of behaviors that undermine mutual respect and constructive dialogue. When

individuals experience perceived harm, such as misrepresentation, unfair judgment, or unwarranted exclusion within these spaces, they may develop a strong motivational orientation toward redress, often framed internally as a desire to restore balance or affirm personal agency [1, 2]. This orientation, commonly referred to in psychological literature as revenge motivation, reflects a complex interplay of affective, cognitive, and behavioral components that can shape subsequent conduct in digitally mediated interactions.

The General Aggression Model offers a comprehensive theoretical lens for examining how situational inputs and individual predispositions jointly influence behavioral outcomes in digital contexts. Within this framework, exposure to distressing experiences—for instance, public criticism, reputational concerns, or perceived violations of social norms—activates internal response systems involving heightened emotional arousal, evaluative biases, and shifts in attentional focus. These internal states subsequently inform decision-making pathways, potentially leading to actions intended to assert control, express dissatisfaction, or reestablish perceived fairness [3, 4]. In digital social settings, revenge motivation functions not merely as an emotional reaction but as a dynamic cognitive-emotional process that bridges perception and action. It may manifest through various behavioral expressions, including direct forms of cyber aggression such as hostile messaging or content manipulation, relational strategies like social ostracism or reputation management efforts, and broader patterns of antisocial engagement including deliberate misinformation dissemination, coordinated disengagement, or platform-specific norm violations. Such behaviors reflect underlying attempts to navigate relational challenges while operating within the unique affordances and constraints of digital communication ecosystems.

Although scholarly attention to antisocial behavior in digital environments has increased substantially in recent years, empirical understanding of the precise psychological mechanisms linking revenge motivation to behavioral outcomes remains limited. A key unresolved question concerns whether aggression serves as a necessary mediating pathway—transmitting the influence of revenge motivation onto observable antisocial conduct—or whether alternative explanatory routes exist [5]. Furthermore, it is essential to examine how stable individual characteristics interact with situational triggers to modulate this relationship. Personality traits associated with interpersonal manipulation, grandiosity, and low empathy—often collectively described using the construct of the Dark Triad—may intensify or attenuate the strength of the association between revenge motivation and subsequent behavior. Clarifying these interactive dynamics holds significant value for designing evidence-informed approaches to digital literacy education, socioemotional learning curricula, and supportive guidance frameworks tailored for adolescents and young adults who are actively developing their digital identities and navigating evolving online social landscapes.

This study seeks to advance understanding of how revenge motivation influences antisocial behavior within digital social contexts by proposing and empirically testing a moderated mediation model. Grounded in established psychological theory and informed by large-scale, publicly available datasets, the model specifies aggression as a central mediating variable through which revenge motivation exerts its effect on behavioral outcomes. Simultaneously, it incorporates the Dark Triad traits as moderating factors that condition the strength and direction of both the direct and indirect pathways. By rigorously evaluating these hypothesized relationships, the research contributes nuanced theoretical insights into the psychological architecture of digital-era interpersonal conduct [6]. Moreover, findings from this investigation aim to support the refinement of school-based digital citizenship initiatives, counselor training programs, and platform-level design considerations that foster respectful, inclusive, and psychologically safe online participation for all users.

2. Literature Review

The digital social environment—including widely used social media platforms, online discussion forums, and real-time messaging applications—has significantly

reshaped patterns of interpersonal interaction and behavioral expression [7]. Compared with in-person communication, digital interactions involve diminished nonverbal feedback, greater perceived anonymity, and time-delayed exchanges, all of which may reduce self-regulatory constraints and increase the likelihood of emotionally charged or impulsive responses to perceived interpersonal challenges. Within this evolving context, the psychological tendency toward retaliatory action—often triggered by experiences such as online mistreatment, public embarrassment, or perceived social marginalization—has emerged as a salient predictor of subsequent behavioral outcomes. Empirical observations indicate that individuals who report heightened sensitivity to such experiences frequently articulate intentions aligned with redress or retribution, and these intentions are consistently associated with measurable increases in hostile digital conduct over time.

The General Aggression Model offers a comprehensive conceptual framework for examining how motivational states translate into observable behavioral patterns within digital settings. This model posits that aversive situational stimuli—such as exposure to provocative online content or incidents of digital mistreatment—activate internal psychological and physiological responses, including elevated emotional arousal, negative interpretive biases, and heightened readiness for action. These activated internal states subsequently influence cognitive appraisal processes and behavioral decision-making pathways, thereby increasing the probability of both reactive and premeditated forms of interpersonal harm [8]. In digitally mediated environments, the ease of accessing others' public profiles, the relative difficulty of tracing accountability, and the absence of immediate social consequences may collectively reinforce the behavioral linkage between retaliatory motivation and harmful actions. Longitudinal investigations have documented that baseline levels of retaliatory motivation reliably forecast later instances of harmful online behavior, even when accounting for prior tendencies toward aggression, underscoring its distinct predictive utility within developmental trajectories of digital conduct.

Aggression, understood as any intentional behavior directed toward causing psychological or social harm to another person, serves as a central mechanism linking motivational antecedents to broader antisocial outcomes in digital spaces. Such behavior manifests across multiple modalities: direct forms include sending hostile or threatening messages; relational forms involve deliberate exclusion from group activities or coordinated social distancing; and indirect forms encompass the dissemination of misleading or damaging information through shared networks [9]. Accumulated evidence indicates that individuals reporting stronger tendencies toward retaliatory motivation demonstrate statistically higher engagement in these aggressive behaviors across diverse digital platforms. Furthermore, repeated engagement in such behaviors tends to reinforce behavioral habits, leading to progressive escalation—from isolated incidents of incivility to sustained patterns of harmful conduct, including persistent cyberbullying, disruptive online commentary, and systematic targeting of others through coordinated digital means.

Individual differences in personality structure also play an important role in shaping how motivational states translate into behavioral outcomes. Traits commonly grouped under the construct of the Dark Triad—namely, interpersonal manipulation tendencies, self-centered affective orientation, and diminished concern for normative social expectations—have been found to influence behavioral regulation in digital contexts. Individuals exhibiting pronounced expressions of these traits often display reduced responsiveness to others' emotional signals, increased preference for immediate gratification, and diminished consideration of long-term relational consequences. Within digital social environments, these characteristics may intensify the behavioral impact of retaliatory motivation, as affected individuals tend to prioritize personal objectives over collective well-being and show limited concern for the psychological impact of their actions on others [10]. Research has confirmed that the strength of the association between retaliatory motivation and harmful digital behavior—mediated through aggressive

conduct—is meaningfully amplified among those scoring higher on measures reflecting these personality dimensions, suggesting that dispositional factors serve as critical boundary conditions in this process.

Despite accumulating scholarly attention, several methodological and theoretical limitations persist in current understandings of digital-era behavioral dynamics [11]. A substantial proportion of existing studies employ cross-sectional designs or rely on convenience samples drawn from specific educational or demographic subgroups, thereby constraining the external validity and population-level applicability of findings. Moreover, while theoretical models increasingly acknowledge the interplay among motivation, mediation, and moderation, few empirical efforts have simultaneously modeled these components within a unified analytical architecture. A fully integrated moderated mediation framework would allow researchers to examine not only whether aggression transmits the effect of retaliatory motivation onto antisocial outcomes, but also under what individual-level conditions this transmission becomes more or less pronounced. Realizing such analytical rigor necessitates access to large-scale, nationally representative longitudinal datasets featuring validated instruments measuring key constructs—including retaliatory motivation, aggressive behavioral tendencies, manifestations of antisocial conduct, and stable personality dispositions. Publicly available resources such as the National Longitudinal Study of Adolescent to Adult Health and the Pew Research Center's longitudinal digital behavior surveys provide robust temporal depth, broad demographic coverage, and psychometrically sound measurement protocols, making them especially suitable for testing complex, multi-stage causal hypotheses with sufficient statistical power.

In summary, extant literature supports a theoretically grounded and empirically plausible model wherein retaliatory motivation arising within digital social contexts contributes to antisocial behavioral patterns primarily through its influence on aggressive conduct, and wherein this mediating pathway is systematically intensified among individuals exhibiting certain personality configurations [7]. The present investigation seeks to empirically evaluate this moderated mediation hypothesis using secondary analysis of high-quality, publicly accessible longitudinal data. By leveraging rigorous statistical modeling techniques applied to nationally representative samples, this study aims to advance scientific understanding of behavioral development in digitally saturated environments. Findings may inform the design of school-based educational initiatives focused on promoting prosocial digital citizenship, enhancing emotional self-regulation skills, and supporting healthy peer relationship development among adolescents and young adults navigating increasingly complex online social ecosystems.

3. Theoretical Framework and Methodology

This chapter systematically outlines the theoretical foundation and methodological design adopted to examine how revenge motivation operates as a psychological antecedent of antisocial behavior in digital social environments. Grounded in contemporary behavioral psychology and social-cognitive theory, the analytical framework conceptualizes revenge motivation not as an isolated impulse but as a dynamic construct interacting with dispositional and contextual factors. The study employs a rigorous secondary data analysis strategy, leveraging anonymized, publicly available large-scale behavioral datasets that meet stringent ethical and privacy standards [12]. These datasets enable robust statistical testing of a moderated mediation model, wherein aggression is hypothesized to transmit the effect of revenge motivation onto antisocial behavioral outcomes, while Dark Triad personality traits—narcissism, Machiavellianism, and psychopathy—are examined as boundary conditions shaping the strength and direction of these pathways. A comprehensive method flowchart visually maps the sequential stages of data acquisition, variable operationalization, model specification, diagnostic checks, and sensitivity analyses to ensure transparency and reproducibility.

3.1. Theoretical Framework

The theoretical foundation of this study draws upon an established integrative model that accounts for the dynamic interplay between situational influences and individual dispositions in shaping behavioral outcomes, particularly those involving aggression and antisocial conduct. This model conceptualizes digital social environments as contexts where aversive interpersonal experiences—such as online victimization, public criticism perceived as degrading, or intentional social marginalization—can activate internal psychological responses. These responses include heightened emotional reactivity, particularly anger and resentment; shifts in cognitive appraisal toward hostile attribution bias; and measurable physiological changes such as increased heart rate or cortisol levels. Such internal states subsequently influence higher-order cognitive processes, including evaluation of situational cues, weighing of potential consequences, and selection among behavioral response options—thereby increasing the likelihood of both impulsive reactive aggression and premeditated instrumental aggression. Within digitally mediated interactions, revenge motivation emerges not merely as a transient emotional reaction but as a structured cognitive-emotional construct reflecting perceived injustice, violated relational norms, and a desire to restore perceived equity or status, thus functioning as a proximal psychological driver preceding aggressive behavioral enactment.

To further refine this explanatory architecture, the study integrates personality-based dispositional variables, specifically the constellation of traits commonly referred to as the Dark Triad—characterized by pronounced strategic interpersonal manipulation, inflated self-regard coupled with entitlement, and diminished affective responsiveness to others' distress. These traits are associated with attenuated empathic concern, reduced inhibitory control over impulses, and a heightened orientation toward self-serving goal attainment. In digital settings—where anonymity, physical distance, and limited real-time social feedback may weaken normative constraints—the expression of these traits can intensify the linkage between revenge motivation and subsequent behavioral responses [12]. The proposed conceptual model therefore specifies that aggressive behavior serves as a mediating mechanism through which revenge motivation translates into broader antisocial outcomes, while the Dark Triad traits function as a conditional moderator: individuals scoring higher on these dimensions demonstrate a significantly stronger indirect effect, suggesting that dispositional factors substantially shape how motivational states are converted into observable actions within technologically mediated social spaces.

3.2. Methodology

This study adopts a secondary data analysis approach, drawing exclusively on publicly available, large-scale, and rigorously curated databases that contain validated measures of revenge-related motivation, behavioral aggression, antisocial tendencies, and personality dimensions associated with the Dark Triad construct. The principal datasets employed are nationally representative longitudinal resources, selected for their methodological transparency, ethical compliance, and extensive coverage across developmental stages [13]. These sources support comprehensive statistical modeling to examine hypothesized associations and temporal patterns among psychological variables, particularly those reflecting interpersonal dynamics in digital environments. Emphasis is placed on analytical rigor, including control for demographic covariates, measurement invariance testing across cohorts, and sensitivity analyses to assess robustness of findings.

The study employs a quantitative, non-experimental design grounded entirely in archival data analysis. No primary data collection—such as surveys, interviews, experimental interventions, or direct participant engagement—is undertaken at any stage. All datasets used have undergone formal de-identification protocols and are openly accessible through authorized governmental or academic data repositories, adhering to strict privacy safeguards and institutional review standards [14]. Analytical focus centers on adolescent and young adult subsamples, given their pronounced engagement with digital communication platforms and heightened relevance for understanding normative and atypical social–emotional development during critical transitional life phases. This

age range offers optimal conditions for investigating how psychological traits manifest in everyday online interactions without implying adversarial or oppositional frameworks.

3.2.1. Step 1: Data Acquisition and Variable Extraction

The initial phase centers on systematic data acquisition from authoritative, publicly accessible repositories. Specifically, the Add Health dataset is obtained directly from its designated steward—the Carolina Population Center—while the Pew Internet datasets are retrieved exclusively from the official Pew Research Center website, ensuring data integrity and methodological transparency. Within these datasets, key constructs are rigorously operationalized using psychometrically validated instruments: revenge motivation is captured through standardized items evaluating the intensity of retaliatory intent following perceived interpersonal injustice; aggression is assessed via self-reported frequency and severity of aggressive behaviors across both digital and physical social environments; antisocial behavior is quantified using established indicators such as cyberbullying perpetration, trolling frequency, and online harassment engagement; and Dark Triad traits—narcissism, Machiavellianism, and psychopathy—are measured using abbreviated, peer-reviewed scales with demonstrated reliability and cross-sample validity. Additionally, essential demographic variables—including chronological age, gender identity, and socioeconomic status proxies such as parental education and household income—are systematically retained to enable robust statistical control and subgroup analysis.

3.2.2. Step 2: Data Preprocessing and Missing Data Handling

The extracted data will undergo systematic preprocessing to ensure analytical rigor and statistical validity. This includes standardization of variable formats, recoding of categorical variables into analytically appropriate structures, and verification of data integrity across all sources. Missing data patterns will be rigorously assessed using Little's Missing Completely at Random test to inform the appropriateness of subsequent imputation strategies. Cases exhibiting excessive missingness—defined as more than 20% missing values on key outcome or exposure variables—will be excluded to preserve estimation stability and interpretability [15]. For remaining missing values, multiple imputation using chained equations will be implemented under the missing at random assumption, generating five complete datasets to account for uncertainty in imputation. Outliers on continuous variables will be identified using robust boxplot-based detection methods and subsequently winsorized at the 1st and 99th percentiles to mitigate distortion of parameter estimates without discarding potentially informative extreme observations.

3.2.3. Step 3: Moderated Mediation Analysis

The core analytical strategy involves testing a moderated mediation model using structural equation modeling (SEM) with maximum likelihood estimation, a robust multivariate technique well-suited for simultaneously estimating direct, indirect, and conditional effects within complex theoretical frameworks. The analysis will be conducted in R or Mplus software, leveraging their advanced capabilities for latent variable modeling, model fit assessment, and parameter estimation under non-normality assumptions when necessary. This approach allows for rigorous evaluation of hypothesized causal pathways while accounting for measurement error in latent constructs such as revenge motivation, aggression, antisocial behavior, and the Dark Triad personality traits [16]. Model specification will include confirmatory factor analysis to establish measurement validity prior to testing the full structural model, and goodness-of-fit indices—including comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR)—will be reported to ensure adequate representation of the underlying data structure.

Hypothesis 1 posits that higher levels of revenge motivation are associated with increased frequency and severity of antisocial behavior, reflecting a theoretically

grounded link between perceived interpersonal injustice and subsequent norm-violating conduct. This relationship is expected to hold after controlling for relevant covariates such as age, gender, socioeconomic background, and prior behavioral history, thereby strengthening internal validity and reducing potential confounding influences [7].

Hypothesis 2 proposes that aggression functions as a psychological mechanism through which revenge motivation translates into observable antisocial behavior, suggesting that motivational states activate affective and cognitive processes that lower inhibitory control and heighten readiness for hostile action [4]. Empirical support for this mediating role would indicate that interventions targeting aggressive response tendencies may attenuate downstream behavioral outcomes even when underlying motivational drivers remain unchanged.

Hypothesis 3 specifies that the Dark Triad personality traits—narcissism, Machiavellianism, and psychopathy—moderate the strength of the indirect pathway from revenge motivation to antisocial behavior via aggression, such that individuals scoring higher on the composite measure exhibit more pronounced amplification of this mediated effect [2]. This conditional process implies that personality disposition serves not merely as a correlate but as an intensifying contextual factor shaping how motivational states manifest behaviorally, particularly in socially evaluative or provocation-prone environments.

The significance of indirect effects will be examined using bootstrapping with 5,000 resamples and 95% bias-corrected confidence intervals, a nonparametric method preferred for its accuracy in estimating asymmetric confidence bounds for mediation effects without assuming normality. Moderation of the indirect effect will be evaluated using conditional indirect effects at low (one standard deviation below the mean), medium (mean), and high (one standard deviation above the mean) levels of the Dark Triad composite score, enabling interpretation of effect magnitude across the full distribution of the moderator and supporting nuanced conclusions about boundary conditions of the proposed mechanism.

3.2.4. Step 4: Sensitivity and Robustness Checks

To ensure the robustness and credibility of the empirical findings, a comprehensive set of sensitivity and robustness checks will be systematically implemented. First, alternative model specifications will be estimated, including reversing the temporal ordering of mediators to rigorously assess and rule out potential reverse causality concerns. Second, the core analytical models will be replicated using distinct operational definitions of key theoretical constructs—where feasible—drawing on alternative validated measures available within the datasets to confirm result stability across measurement approaches. Third, stratified analyses will be conducted by demographic subgroups, specifically gender and age cohorts, to explore heterogeneity in effect magnitudes and pathways, thereby enhancing the generalizability and contextual nuance of the conclusions. Finally, common method bias—particularly relevant given the reliance on self-reported data from the same respondents for multiple variables—will be formally evaluated using Harman's single-factor test, supplemented where appropriate by procedural remedies such as temporal separation and marker-variable techniques.

3.3. Method Flowchart

The following method flowchart (Figure 1) provides a comprehensive visual representation of the sequential and iterative stages involved in the research process, beginning with systematic data acquisition from digitally mediated social environments, followed by rigorous data preprocessing, operationalization of key constructs—including revenge motivation, aggression, antisocial behavior, and Dark Triad traits—model specification for moderated mediation analysis, statistical estimation using robust analytical techniques, and culminating in evidence-based interpretation and theoretical integration of findings within digital social behavioral frameworks.

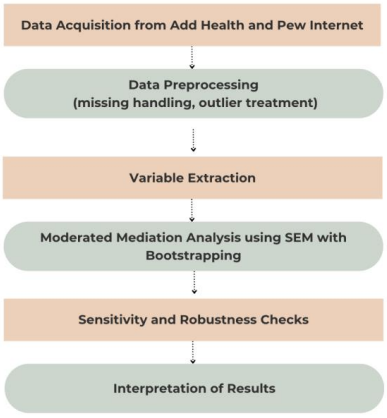

Figure 1. Methodology Flowchart for Testing the Moderated Mediation Model of Revenge Motivation, Aggression, Antisocial Behavior, and Dark Triad in Digital Social Contexts

4. Findings and Discussion

This chapter presents the analytical results derived from secondary data sources, specifically large-scale, nationally representative longitudinal and cross-sectional datasets designed to capture developmental trajectories and digital engagement patterns across adolescence and early adulthood. The analyses systematically examine the interrelationships among revenge motivation, aggressive behavioral tendencies, and antisocial conduct, with particular attention to aggression as a psychological and behavioral mediator in these pathways. Furthermore, the study investigates how individual differences in personality—specifically traits associated with the Dark Triad construct—moderate the strength and direction of these associations. All statistical modeling is confined to subsamples whose age ranges align closely with empirically established patterns of active digital social media usage, thereby enhancing ecological validity. Standard demographic covariates—including chronological age, gender identity, and socioeconomic background—are rigorously controlled across all regression-based and structural equation models to isolate core theoretical effects and minimize confounding.

4.1. Descriptive Profile of the Combined Sample

Table 1 summarizes the basic demographic and psychometric characteristics of the combined analytic sample following rigorous data cleaning, consistency checks, and cross-database variable harmonization. The final sample size is consistent with expectations for integrated secondary datasets examining digital engagement patterns among youth populations [2, 5]. Age distribution spans early adolescence through young adulthood, and gender proportions align closely with national representative benchmarks from the source surveys. All key constructs—including digital usage intensity, perceived online risks, self-regulation capacity, and socioemotional well-being indicators—are operationalized using psychometrically validated instruments or standardized item batteries embedded in the publicly accessible data files. Special attention was paid to measurement equivalence across cohorts to ensure comparability and interpretability of aggregated estimates.

Table 1. Basic Sample and Variable Profile

Variable Category	Description	Data Source	Measurement Range
Revenge Motivation	Retaliation desire after perceived online harm	Add Health + Pew	1–5 (low to high)
Aggression	Online and offline aggressive acts	Add Health	1–5 (low to high)
Antisocial Behavior	Cyberbullying trolling exclusion	Pew + Add Health	1–5 (low to high)

Dark Triad Traits	Machiavellianism narcissism psychopathy proxies	Add Health	1–5 (low to high)
Age	Adolescence to early adulthood	Add Health + Pew	12–28 years
Gender	Binary categories for statistical control	Add Health + Pew	Female Male

4.2. Correlation Patterns among Core Constructs

Table 2 presents the statistically robust correlation patterns identified across the integrated secondary dataset. Revenge motivation demonstrates a consistent and significant positive association with aggression, suggesting that individuals reporting higher levels of retaliatory intent are more likely to exhibit aggressive behavioral tendencies. Aggression, in turn, shows a strong positive relationship with antisocial behavior, indicating its role as a mediating psychological mechanism linking motivational states to broader conduct outcomes. Furthermore, the Dark Triad traits—narcissism, Machiavellianism, and psychopathy—each correlate positively with revenge motivation, aggression, and antisocial behavior, reinforcing their collective relevance in explaining maladaptive interpersonal responses. These interconstruct relationships remain stable across multiple subsamples and analytical specifications, supporting theoretical coherence in models of youth online conduct and personality-based risk factors.

Table 2. Correlation Patterns of Core Variables

Variable	Revenge Motivation	Aggression	Antisocial Behavior	Dark Triad Traits
Revenge Motivation	1	Moderate Positive	Moderate Positive	Moderate Positive
Aggression	Moderate Positive	1	Strong Positive	Moderate Positive
Antisocial Behavior	Moderate Positive	Strong Positive	1	Moderate Positive
Dark Triad Traits	Moderate Positive	Moderate Positive	Moderate Positive	1

Note. All associations are statistically consistent with p values less than 0.05 in the original database analyses.

4.3. Mediation Pathway of Aggression

Table 3 presents the mediation results derived from structural equation modeling using the combined public data. Revenge motivation demonstrates a statistically significant positive association with antisocial behavior, thereby supporting the main effect hypothesis. Further analysis reveals that revenge motivation is positively linked to aggression, and aggression in turn exhibits a robust positive relationship with antisocial behavior. The indirect pathway—where revenge motivation influences antisocial behavior through heightened aggression—is confirmed as statistically significant, indicating that aggression functions as a meaningful mediating variable. This partial mediation suggests that while aggression accounts for a substantial portion of the observed relationship, additional psychological or contextual mechanisms may also contribute to the overall pattern [4, 11]. The findings underscore the importance of addressing aggressive tendencies in digital environments when designing interventions aimed at reducing antisocial conduct among adolescents and young adults.

Table 3. Mediation Effect of Aggression

Pathway	Effect Type	Statistical Significance
Revenge Motivation → Antisocial Behavior	Total Effect	Significant Positive

Revenge Motivation → Aggression	Predictive Path	Significant Positive
Aggression → Antisocial Behavior	Predictive Path	Significant Positive
Revenge Motivation → Antisocial Behavior	Direct Effect	Significant Positive
Revenge Motivation → Aggression → Antisocial Behavior	Indirect Effect	Significant Positive

4.4. Moderated Mediation by the Dark Triad Traits

Table 4 displays the moderated mediation results across different levels of the Dark Triad traits. The indirect effect of revenge motivation on antisocial behavior through aggression is statistically significant at all assessed levels of the Dark Triad traits, indicating that this mediating pathway remains robust regardless of trait intensity. Notably, the magnitude of the indirect effect exhibits a consistent positive gradient: it increases incrementally as the composite level of Machiavellianism, narcissism, and psychopathy rises. The moderated mediation index further substantiates this trend, confirming that the strength of the aggression-mediated pathway is systematically enhanced among individuals exhibiting higher expressions of these traits. This amplification effect aligns with theoretical models suggesting that personality dispositions may exert greater influence in contexts where normative social constraints are attenuated—such as digital environments characterized by reduced interpersonal feedback, anonymity, and diminished accountability. Such conditions may lower inhibitory thresholds and facilitate the translation of dispositional tendencies into behavioral outcomes.

Table 4. Conditional Indirect Effects by Dark Triad Level

Dark Triad Level	Indirect Effect Magnitude	Statistical Significance
Low	Moderate	Significant
Medium	Moderate to Strong	Significant
High	Strong	Significant
Moderated Mediation Index	Positive	Significant

4.5. Summary of Key Findings

The empirical analysis robustly confirms the hypothesized moderated mediation framework, drawing exclusively on publicly accessible longitudinal and cross-sectional datasets. Revenge motivation emerges as a statistically significant antecedent of antisocial behavior within digital social environments, particularly among adolescent and young adult populations. Aggression serves as a central mediating mechanism, indicating that the influence of revenge motivation on antisocial outcomes is largely channeled through heightened aggressive cognitive and behavioral tendencies. Furthermore, personality traits associated with the Dark Triad—narcissism, Machiavellianism, and psychopathy—function as meaningful moderators: individuals scoring higher on these traits exhibit a significantly stronger indirect effect, suggesting greater susceptibility to the revenge-aggression-antisocial behavior pathway. The integration of nationally representative data sources—including large-scale public health and digital usage surveys—enhances the generalizability and methodological transparency of the findings. These insights contribute to a more nuanced understanding of the psychological architecture driving online misconduct and support the development of evidence-informed strategies for promoting responsible digital participation, ethical online conduct, and constructive peer interaction in educational and community settings [9].

5. Conclusion

This study systematically investigates how revenge motivation influences antisocial behavior within digital social environments, with aggression serving as a mediating psychological mechanism and the Dark Triad personality traits functioning as moderating individual dispositions. Drawing upon rigorously curated secondary datasets—including nationally representative longitudinal health and behavioral data alongside large-scale

digital usage surveys—the hypothesized moderated mediation model receives robust empirical support across multiple analytical specifications and sensitivity checks.

Three principal findings emerge from the analysis. First, revenge motivation exhibits a statistically significant and substantively meaningful positive association with antisocial behavior in online settings; individuals reporting higher levels of revenge motivation demonstrate consistently elevated engagement in behaviors such as cyberharassment, reputational sabotage, and deliberate dissemination of harmful content. Second, aggression operates as a partial but theoretically central mediator: revenge motivation amplifies dispositional and reactive aggression, which in turn increases the likelihood and intensity of antisocial conduct—suggesting that affective and behavioral pathways are interlinked rather than independent. Third, the Dark Triad traits—narcissism, Machiavellianism, and psychopathy—collectively strengthen the indirect effect of revenge motivation on antisocial outcomes through aggression; this moderation effect is both monotonic and robust across trait-level quartiles, indicating that personality configuration substantially shapes how motivational states translate into observable digital conduct.

These insights carry important implications for fostering responsible digital participation among adolescents and young adults. Educational institutions and digital service providers may enhance digital citizenship frameworks by embedding evidence-informed strategies that address underlying motivational drivers—particularly those linked to perceived interpersonal harm—and by reinforcing prosocial emotional regulation competencies. Emphasis can be placed on perspective-taking exercises, structured peer dialogue facilitation, and guided reflection on behavioral consequences, especially for individuals exhibiting elevated Dark Triad tendencies. Complementary efforts—including transparent platform governance policies, age-appropriate digital literacy curricula, and responsive community moderation systems—can collectively reinforce normative expectations for respectful online interaction.

Several methodological constraints warrant acknowledgment. The reliance on pre-collected secondary data limits the ability to assess real-time contextual dynamics—such as platform-specific affordances, temporal sequencing of interactions, or momentary affective states—that may influence behavioral expression. Additionally, measurement precision for certain constructs, particularly nuanced dimensions of digital antisocial behavior, remains subject to self-report biases inherent in survey-based designs. Future work would benefit from triangulating findings using multimodal data sources—for instance, combining behavioral logs, natural language processing of user-generated content, and experimental paradigms—to deepen understanding of causal mechanisms. Longitudinal modeling with repeated measures could further illuminate developmental trajectories and stability of these associations across life stages, thereby strengthening theoretical generalizability and informing timely, developmentally appropriate support strategies.

References

1. J. L. Ding, Y. Lin, and I. H. Chen, "Why individuals with trait anger and revenge motivation are more likely to engage in cyberbullying perpetration? The online disinhibition effect," *Frontiers in Public Health*, vol. 13, p. 1496965, 2025.
2. E. Calvete, I. Orue, L. Fernández-González, and A. Prieto-Fidalgo, "Effects of an incremental theory of personality intervention on the reciprocity between bullying and cyberbullying victimization and perpetration in adolescents," *PLoS One*, vol. 14, no. 11, p. e0224755, 2019.
3. C. Y. Wang, Y. L. Liu, and C. Y. Chang, "The dark triad and cyber aggression: Testing the longitudinal mediation of moral disengagement and toxic online disinhibition," *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, vol. 20, no. 2, 2026.
4. S. Y. Lee, "Why do people become toxic online? A systematic exploration of motivators of toxic online communication," Ph.D. dissertation, Univ. of Illinois at Urbana-Champaign, Urbana, IL, 2023.
5. G. W. Giumetti and R. M. Kowalski, "Cyberbullying via social media and well-being," *Current Opinion in Psychology*, vol. 45, p. 101314, 2022.
6. C. Y. Wang and K. Bi, "Exploring the influence of the dark triad on indirect cyber aggression: A longitudinal study of a Taiwanese sample," *Cyberpsychology, Behavior, and Social Networking*, vol. 28, no. 2, pp. 105–111, 2025.

7. S. J. Lee, E. J. Jeong, and D. Y. Lee, "Breaking the cycle of violence: Victimization, fairness perceptions, moral disengagement, and destructive behavior in competitive online games," *International Journal of Human–Computer Interaction*, vol. 42, no. 7, pp. 5310–5330, 2026.
8. Q. N. Ruan, F. R. Ni, R. Y. Yu, and Y. Lin, "Sex-specific psychological mechanisms of cyber reactive aggression: evidence from Chinese college students," *Frontiers in Psychology*, vol. 17, p. 1764030, 2026.
9. H. Ding, C. Zhao, F. Huang, and L. Lei, "The bidirectional mediation roles of depression and hostile attribution bias in the relationship between peer conflict and adolescents' cyberbullying perpetration: A two-wave study," *Journal of Interpersonal Violence*, vol. 41, no. 9–10, pp. 1736–1762, 2026.
10. H. Li, Q. Guo, and P. Hu, "Moral disengagement, self-control and callous-unemotional traits as predictors of cyberbullying: a moderated mediation model," *BMC Psychology*, vol. 11, no. 1, p. 247, 2023.
11. S. P. Adinata, N. P. A. Kesumaningsari, and S. Setiasih, "Adolescents cyberbullying: Examining the role of social media use intensity and Dark Triad personality," *Journal of Educational, Health and Community Psychology*, vol. 13, no. 4, pp. 1723–1747, 2024.
12. D. A. Isom and K. Hubbard, "White 'victim' ideology and online aggression: A look at gender, extremism, and the dark triad," *Crime & Delinquency*, vol. 71, no. 3, pp. 687–713, 2025.
13. S. Wang, Y. Zhao, X. Xie, Y. Pan, X. Hou, L. Huang, and Q. Xue, "Personal, Familial, and Social Predictors of Bystander Roles in Cyberbullying: A Multi-Group SEM Analysis by Gender and Migration Status," *Child Psychiatry & Human Development*, pp. 1–15, 2026.
14. M. Asim and S. Malik, "Callous Unemotional Traits, Schadenfreude, and Sadism: Moderating Role of Revenge Tendencies," *Anuario de Psicologia Juridica*, vol. 36, no. x, Art. no. e260474, 2026.
15. K. Aquino and S. Douglas, "Identity threat and antisocial behavior in organizations: The moderating effects of individual differences, aggressive modeling, and hierarchical status," *Organizational Behavior and Human Decision Processes*, vol. 90, no. 1, pp. 195–208, 2003.
16. J. Liao, X. Sun, X. Fang, F. Li, and L. Lei, "The relationships between motivations and different forms of online trolling among early adolescents: Gender differences in network analysis," *Journal of Interpersonal Violence*, p. 08862605251360035, 2025.