

**The Role of Online Disinhibition and Anonymity in Predicting
Cyberbullying Among University Students**

Shehzad Khattak

shehzadkhattak5544@gmail.com

Department of Psychology, Abdul Wali Khan University Mardan, Khyber Pakhtunkhwa, Pakistan

Sobia Shah

sobiashah2030@gmail.com

Department of Psychology, Abdul Wali Khan University Mardan, Khyber Pakhtunkhwa, Pakistan

Aneesa

aneesarawail2040@gmail.com

Department of Psychology, Abdul Wali Khan University Mardan, Khyber Pakhtunkhwa, Pakistan

Muskan Sayyar

muskansayyar22@gmail.com

Department of Psychology, Abdul Wali Khan University Mardan, Khyber Pakhtunkhwa, Pakistan

Corresponding Author: Shehzad Khattak shehzadkhattak5544@gmail.com

Received: 11-02-2026

Revised: 25-02-2026

Accepted: 13-03-2026

Published: 28-03-2026

ABSTRACT

Cyberbullying has emerged as a significant psychosocial and educational concern among university students in South Asia, yet the psychological mechanisms underlying this behavior remain under-investigated in this cultural context. The present study examined the role of online disinhibition and perceived anonymity in predicting cyberbullying perpetration among university students in Khyber Pakhtunkhwa (KPK), Pakistan. A quantitative, cross-sectional survey design was employed with a convenience sample of 300 university students (55% male, 45% female; M age range: 22–25 years) drawn from public and private universities in KPK. Participants completed the Online Disinhibition Scale (ODS; Udris, 2014), the Perceived Anonymity Scale (Barlett et al., 2016), and the Cyberbullying Inventory for University Students (Tanrikulu & Erdur-Baker, 2021). Data were analyzed using descriptive statistics, reliability analysis, Pearson product-moment correlations, and multiple regression analysis. Results revealed significant positive associations between online disinhibition and cyberbullying ($r = .58, p < .01$), between perceived anonymity and cyberbullying ($r = .49, p < .01$), and between online disinhibition and perceived anonymity ($r = .46, p < .01$). Multiple regression analysis indicated that online disinhibition ($\beta = .43, p < .001$) and perceived anonymity ($\beta = .29, p < .001$) were both significant independent predictors of cyberbullying. The overall model was statistically significant, $F(2, 297) = 102.85, p < .001$, with the two predictors jointly accounting for 41% of the variance in cyberbullying behavior ($R^2 = .41$). Findings are interpreted within the framework of Suler's (2004) Online Disinhibition Effect Theory and Barlett and Gentile's (2012) general theory of online aggression, and implications for digital ethics education, university policy, and counseling services are discussed.

Keywords: online disinhibition, perceived anonymity, cyberbullying, university students, online aggression, deindividuation, digital ethics

INTRODUCTION

The rapid proliferation of digital communication technologies has fundamentally transformed the social, academic, and psychological landscape of contemporary university life. Social media platforms, instant messaging applications, online gaming environments, and web-based collaborative tools are now integral to how university students communicate, learn, and form social identities (Faucher et al., 2014). While these technologies afford significant opportunities for connection and knowledge sharing, they simultaneously create new ecological contexts in which interpersonal harm can occur with unprecedented reach, speed, and persistence. Among the most consequential of these harms is cyberbullying—broadly defined as the use of electronic communication technologies to intentionally and repeatedly harm or harass another individual (Smith et al., 2008).

The prevalence and psychosocial impact of cyberbullying among university students have attracted growing scholarly attention over the past two decades. Research conducted across multiple countries indicates that between 10% and 40% of university students report cyberbullying victimization, with variation attributable to definitional, methodological, and cultural differences (Faucher et al., 2014; Mishna et al., 2020; Schenk & Fremouw, 2012). The documented consequences of cyberbullying for victims are wide-ranging and clinically significant, encompassing depression, anxiety, diminished academic performance, social withdrawal, and, in severe cases, suicidal ideation (Kowalski et al., 2014). Perpetrators are likewise not immune to negative outcomes, with cyberbullying perpetration associated with disciplinary consequences, social ostracism, and legal liability.

Two theoretically salient psychological mechanisms have been consistently implicated in cyberbullying perpetration: online disinhibition and perceived anonymity. Online disinhibition, as formally conceptualized by Suler (2004), refers to the tendency to express oneself online in ways that depart substantially from one's offline behavioral norms—including increased impulsivity, aggression, and disregard for social constraints. Suler proposed six underlying mechanisms: dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination, and minimization of authority. Of particular relevance to harmful online behavior is what Suler termed toxic disinhibition, characterized by hostile communication, threats, and aggression, as distinct from benign disinhibition, which encompasses prosocial self-disclosure and emotional openness.

Perceived anonymity constitutes a closely related yet conceptually distinct construct. In the online environment, anonymity is not merely a technical status but a psychological experience—the subjective sense that one's identity is concealed or unrecognizable to others (Christopherson, 2007). Drawing on Festinger et al.'s (1952) deindividuation theory, which posits that anonymity reduces self-awareness and normative self-regulation, researchers have demonstrated that individuals who feel anonymous in digital settings are more likely to engage in antisocial behaviors including online aggression (Barlett & Gentile, 2012; Barlett, 2015). The mechanism is theorized to operate via reduced fear of punishment and social accountability: those who believe they cannot be identified feel free to act in ways they would otherwise inhibit.

Although both constructs have been studied as predictors of cyberbullying, the majority of published research has examined them in isolation, limiting comprehensive understanding of their joint contribution (Kowalski et al., 2014; Wright, 2014). Furthermore, the existing evidence base is predominantly derived from North American, European, and East Asian contexts, creating a significant geographical gap with respect to South Asian populations (Gorzig & Machackova, 2015). In Pakistan—and specifically in Khyber Pakhtunkhwa (KPK), a province experiencing rapid growth in higher education enrollment and internet connectivity—there is a near-complete absence of empirical research on cyberbullying and its psychological determinants (Pakistan Telecommunication Authority, 2023). Given the distinct cultural

context of KPK, characterized by collectivist social structures, gender-differentiated norms of communication, and variable digital literacy, there are compelling theoretical reasons to examine whether the psychological mechanisms of online disinhibition and perceived anonymity operate in analogous ways to those documented in Western samples.

The present study addresses these gaps by simultaneously examining the independent and joint contributions of online disinhibition and perceived anonymity to cyberbullying perpetration among university students in KPK. Grounded in Suler's (2004) Online Disinhibition Effect Theory and Barlett and Gentile's (2012) general theory of online aggression, the study aims to generate culturally situated empirical evidence that can inform both theoretical understanding and practical intervention in the Pakistani higher education context.

Objectives

The present study pursued the following objectives: (a) to determine the nature and magnitude of the association between online disinhibition and cyberbullying among KPK university students; (b) to examine the relationship between perceived anonymity and cyberbullying; (c) to investigate the association between perceived anonymity and online disinhibition; and (d) to assess the extent to which online disinhibition and perceived anonymity, considered simultaneously, predict cyberbullying perpetration.

Hypotheses

H1: There will be a significant positive relationship between online disinhibition and cyberbullying among university students.

H2: There will be a significant positive relationship between perceived anonymity and cyberbullying among university students.

H3: There will be a significant positive relationship between perceived anonymity and online disinhibition among university students.

H4: Online disinhibition and perceived anonymity will together significantly predict cyberbullying perpetration among university students.

LITERATURE REVIEW

Online Disinhibition: Conceptualization and Empirical Evidence

Suler's (2004) foundational theorization of the online disinhibition effect remains the most widely cited framework for understanding how digital communication environments alter behavioral self-regulation. The core claim is that online contexts systematically attenuate the psychological mechanisms that normally constrain behavior in face-to-face interactions, leading individuals to act in ways that diverge—sometimes dramatically—from their offline behavioral repertoire. Suler identified six interacting factors that produce this effect: dissociative anonymity (the subjective sense that one's online identity is separable from one's real identity); invisibility (the absence of bodily presence and social observation); asynchronicity (the temporal decoupling of online interaction from immediate social consequence); solipsistic introjection (the projection of imagined attributes onto interaction partners); dissociative imagination (the perception of the online world as a rule-governed reality distinct from offline life); and minimization of authority (the erosion of hierarchical social structures that regulate behavior offline).

The distinction between benign and toxic forms of disinhibition is theoretically and practically significant. Benign disinhibition facilitates prosocial processes such as emotional openness, help-seeking, and self-disclosure, and has been productively applied in therapeutic and support-group contexts (Suler, 2004). Toxic disinhibition, by contrast, manifests as aggression, hostility, threats, and antisocial behavior, and represents the psychologically most relevant dimension for understanding cyberbullying (Lapidot-Lefler & Barak, 2012). Empirically, Lapidot-Lefler and Barak (2012) demonstrated through experimental manipulation of anonymity and invisibility conditions that each structural feature of online environments independently elevated hostile disinhibited communication, providing causal evidence for Suler's theoretical mechanisms. Udris (2014) subsequently developed and validated the Online Disinhibition Scale for an adolescent Japanese population, confirming its significant positive association with cyberbullying perpetration. Wright (2014) further demonstrated that online disinhibition partially mediated the relationship between offline social difficulties and online aggressive behavior, establishing disinhibition as a psychological pathway through which pre-existing vulnerabilities translate into harmful online conduct. Most recently, Schacter et al. (2023) found that online disinhibition uniquely predicted cyberbullying perpetration among university students after statistically controlling for trait aggression, social anxiety, and demographic variables.

Perceived Anonymity and Online Aggression

Theoretical accounts of anonymity's behavioral effects are grounded in Festinger et al.'s (1952) deindividuation framework and its subsequent elaborations. Deindividuation theory proposes that conditions diminishing individuality and personal accountability—including physical anonymity, crowd membership, and darkness—reduce self-regulatory inhibitions and increase the likelihood of impulsive, antisocial behavior. Zimbardo's (1969) experimental demonstrations provided initial empirical support, showing that anonymized participants administered higher levels of shock to confederates than identified participants. The application of this framework to digital contexts was systematically theorized by Barlett and Gentile (2012), whose general theory of online aggression posits that perceived anonymity reduces expected punishment, lowers accountability concerns, and thereby increases the probability of aggressive online behavior. This model has received prospective empirical support from Barlett (2015), who demonstrated longitudinally that increases in perceived anonymity temporally preceded increases in cyberbullying perpetration, providing evidence consistent with a causal account. Christopherson (2007) further documented the dual-edged nature of online anonymity, identifying its capacity to facilitate honest self-disclosure among stigmatized populations alongside its documented tendency to enable aggression and deception.

The Relationship Between Online Disinhibition and Perceived Anonymity

Suler (2004) explicitly positioned anonymity as a constitutive mechanism of online disinhibition rather than a merely correlated variable, arguing that the subjective experience of not being recognizable online is psychologically fundamental to the disinhibitory process. This theoretical claim implies that anonymity and disinhibition should be strongly positively correlated, and that anonymity-mediated disinhibition may represent a psychologically coherent pathway to cyberbullying. Lapidot-Lefler and Barak (2012) found empirical support for this relationship, demonstrating that anonymity and invisibility conditions each independently elevated toxic disinhibition, with partially additive effects, consistent with the notion that they represent related but partially separable influences on online behavior. Christopherson (2007) similarly observed that the de-identification of online interactions was associated with increased expression of socially disapproved opinions and behaviors, linking perceived anonymity to a disinhibited communication style. Taken together, these findings suggest that perceived anonymity and online disinhibition are theoretically and empirically interconnected, and that their joint examination within a single predictive model offers a more comprehensive account of cyberbullying risk than either variable alone.

Cyberbullying Among University Students

Cyberbullying is typically defined by three core features: intentionality, repetition, and the digital medium (Smith et al., 2008). However, definitional debates persist around the concept of repetition in online contexts, where a single posting may produce repeated harm through algorithmic amplification or sharing, and around the operationalization of power imbalance in settings where technical capability or social network size may substitute for physical power (Sticca & Perren, 2013). Common forms include harassment, flaming, denigration, impersonation, outing, exclusion, and cyberstalking, each leveraging distinct affordances of digital communication—speed, reach, permanence, and anonymity—to cause harm (Willard, 2007).

Among university students, prevalence estimates across countries range from approximately 9% to 40% for perpetration and victimization respectively (Faucher et al., 2014; Schenk & Fremouw, 2012). University students represent a distinctive risk population owing to their extensive online engagement, reduced institutional oversight compared to secondary school settings, and heightened social pressure around identity formation and peer status (Faucher et al., 2014). Psychosocial consequences of cyberbullying victimization in this population include depression, anxiety, academic disengagement, and social withdrawal (Kowalski et al., 2014; Mishna et al., 2020). Research on cyberbullying in Pakistan remains limited, and no published study has systematically examined the psychological predictors of cyberbullying perpetration among university students in KPK, constituting the primary gap the present study addresses.

Theoretical Framework

The present study is grounded in Suler's (2004) Online Disinhibition Effect Theory, which provides a comprehensive account of how structural features of online environments—particularly anonymity and invisibility—attenuate the psychological restraints that normally govern interpersonal behavior. Within this framework, cyberbullying is conceptualized as a manifestation of toxic disinhibition: the harmful behavioral consequence of reduced self-monitoring, diminished empathy, and loosened normative constraints in the online environment. The study also draws on Barlett and Gentile's (2012) general theory of online aggression, which positions perceived anonymity as a proximal predictor of online aggressive behavior via its effects on accountability expectations, providing a theoretically precise account of the mechanism through which anonymity increases cyberbullying risk. Together, these frameworks generate clear, directional predictions regarding the associations among online disinhibition, perceived anonymity, and cyberbullying, forming the theoretical basis for the hypotheses tested in the present study.

METHODOLOGY

Research Design

A quantitative cross-sectional survey design was employed. This design was selected as appropriate for the study's primary objective of characterizing the nature, direction, and magnitude of associations among online disinhibition, perceived anonymity, and cyberbullying within a defined population at a single time point. No variable was experimentally manipulated; all variables were measured through standardized self-report instruments. While a cross-sectional design cannot establish causal directionality, it permits the computation of multiple regression coefficients that index the unique predictive contributions of each independent variable when the other is statistically controlled—an analytic approach well suited to the study's confirmatory hypotheses.

Participants and Sampling

The target population comprised currently enrolled undergraduate and postgraduate students at public and private universities in Khyber Pakhtunkhwa, Pakistan, aged 18 to 30 years. A total of 300 university students participated in the study ($n = 165$ male, 55.0%; $n = 135$ female, 45.0%). Age was distributed as follows: 18–21 years ($n = 118$, 39.3%), 22–25 years ($n = 142$, 47.3%), and 26–30 years ($n = 40$, 13.4%). Participants were drawn from public ($n = 190$, 63.3%) and private ($n = 110$, 36.7%) universities. Most were undergraduates ($n = 210$, 70.0%), with the remainder enrolled in graduate or postgraduate programs ($n = 90$, 30.0%). A non-probability convenience sampling strategy was employed, with participants selected on the basis of availability, accessibility, and willingness to participate. This approach was adopted given logistical constraints on probability sampling across the geographically dispersed university system of KPK. Demographic characteristics are detailed in Table 1.

Measures

Online Disinhibition Scale (ODS; Udris, 2014). Online disinhibition was assessed using the 11-item Online Disinhibition Scale developed and validated by Udris (2014) for use with students and young adults. Items assess the degree to which respondents feel free, impulsive, or uninhibited in online communication compared to face-to-face interaction. Responses are recorded on a Likert-type scale, with higher total scores indicating greater online disinhibition. In the present sample, the ODS demonstrated good internal consistency ($\alpha = .84$), consistent with values reported in previous applications of the scale.

Perceived Anonymity Scale (Barlett et al., 2016). Perceived anonymity was measured using the 6-item scale developed by Barlett, Gentile, and Chew (2016), which assesses the subjective sense of identity concealment and non-recognizability in online interactions. Items are rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree), with higher scores indicating greater perceived anonymity. The scale demonstrated acceptable internal consistency in the present sample ($\alpha = .79$).

Cyberbullying Inventory for University Students (Tanrikulu & Erdur-Baker, 2021). Cyberbullying perpetration was assessed using the 20-item Cyberbullying Inventory developed and psychometrically validated for use with university populations by Tanrikulu and Erdur-Baker (2021). The scale assesses the frequency of a range of cyberbullying behaviors, including harassment, threatening messages, impersonation, denigration, and spreading harmful rumors, rated on a Likert-type frequency scale (e.g., never to very often). Higher scores reflect greater involvement in cyberbullying behavior. The inventory demonstrated excellent internal consistency in the present sample ($\alpha = .88$).

Procedure

Ethical approval for the study was obtained from the relevant institutional authority at Abdul Wali Khan University Mardan prior to data collection. Participants were recruited through in-person contact at participating universities. Prior to questionnaire administration, the purpose and voluntary nature of the study were explained, and written informed consent was obtained from all participants. Participants were assured of the confidentiality of their responses and were informed that no personally identifying information would be collected or retained. Questionnaires were administered in paper format in private settings to minimize social desirability effects. Participation was estimated to require approximately 20 to 30 minutes. Participants were reminded that they could decline to answer any item or withdraw from the study at any point without consequence.

Data Analysis

Data were analyzed using SPSS Version 26. Prior to primary analyses, data were screened for missing values and incomplete responses; no participant was excluded for this reason. Descriptive statistics (means, standard deviations, ranges) and reliability coefficients (Cronbach's alpha) were computed for all study measures. Pearson product-moment correlation coefficients were calculated to examine bivariate associations among online disinhibition, perceived anonymity, and cyberbullying. A simultaneous multiple regression analysis was then conducted with cyberbullying as the criterion variable and online disinhibition and perceived anonymity as simultaneous predictors, yielding standardized (β) and unstandardized (B) regression coefficients, standard errors, t-values, and model-level fit statistics (R^2 , adjusted R^2 , F-ratio). Assumptions of normality, homoscedasticity, and absence of multicollinearity were verified prior to regression analysis.

RESULTS

Demographic Characteristics

A total of 300 university students participated in the study. The majority were male (55.0%) and enrolled at public universities (63.3%). The modal age group was 22–25 years (47.3%), and most participants were undergraduates (70.0%). Full demographic characteristics are presented in Table 1.

Table 1

Demographic Characteristics of the Sample (N = 300)

Variable	Category	<i>f</i>	%
Gender	Male	165	55.0
	Female	135	45.0
Age	18–21 years	118	39.3
	22–25 years	142	47.3
	26–30 years	40	13.4
University Type	Public	190	63.3
	Private	110	36.7
Academic Level	Undergraduate	210	70.0
	Graduate/Postgraduate	90	30.0

Note. *f* = frequency. Percentages may not sum exactly to 100.0 due to rounding.

Reliability Analysis

Internal consistency estimates for all three study measures were computed using Cronbach's alpha. Results are presented in Table 2. All scales demonstrated acceptable to excellent reliability, with alpha values ranging from .79 (Perceived Anonymity Scale) to .88 (Cyberbullying Inventory), consistent with established benchmarks for psychological research (Nunnally, 1978).

Table 2

Reliability Analysis of Study Measures (N = 300)

Scale	No. of Items	Cronbach's α
Online Disinhibition Scale (ODS; Udris, 2014)	11	.84
Perceived Anonymity Scale (Barlett et al., 2016)	6	.79
Cyberbullying Inventory (Tanrikulu & Erdur-Baker, 2021)	20	.88

Note. α = Cronbach's alpha. Values $\geq .70$ are conventionally regarded as acceptable for research purposes.

Descriptive Statistics

Descriptive statistics for the three primary study variables are presented in Table 3. Participants reported moderate mean levels of online disinhibition ($M = 35.42$, $SD = 7.86$) and perceived anonymity ($M = 20.18$, $SD = 5.21$), alongside moderate mean cyberbullying perpetration scores ($M = 42.76$, $SD = 10.34$). The ranges observed for all three variables indicated adequate variability to support correlational and regression analyses.

Table 3

Descriptive Statistics for Primary Study Variables (N = 300)

Variable	N	M	SD	Min	Max
Online Disinhibition	300	35.42	7.86	15	55
Perceived Anonymity	300	20.18	5.21	7	30
Cyberbullying	300	42.76	10.34	20	78

Note. M = mean; SD = standard deviation.

Pearson Correlation Analysis

Bivariate Pearson correlations among the three primary study variables are presented in Table 4. Online disinhibition was significantly and positively correlated with cyberbullying ($r = .58$, $p < .01$), supporting H1 and indicating that students reporting greater online disinhibition also reported significantly higher levels of cyberbullying perpetration. Perceived anonymity was also significantly and positively associated

with cyberbullying ($r = .49, p < .01$), supporting H2. Furthermore, online disinhibition and perceived anonymity were significantly positively correlated ($r = .46, p < .01$), supporting H3 and demonstrating that students who felt more anonymous online also reported greater disinhibition in their online behavior. All correlations were in the theoretically predicted directions.

Table 4

Pearson Correlations Among Study Variables (N = 300)

Variable	1	2	3
1. Online Disinhibition	—		
2. Perceived Anonymity	.46**	—	
3. Cyberbullying	.58**	.49**	—

Note. ** $p < .01$ (two-tailed).

Multiple Regression Analysis

To test H4, a simultaneous multiple regression analysis was conducted with cyberbullying as the outcome variable and online disinhibition and perceived anonymity as predictors. Results are presented in Table 5. The overall regression model was statistically significant, $F(2, 297) = 102.85, p < .001, R^2 = .41$, adjusted $R^2 = .40$, indicating that online disinhibition and perceived anonymity jointly accounted for 41% of the variance in cyberbullying perpetration.

Online disinhibition emerged as the stronger independent predictor of cyberbullying ($\beta = .43, B = 0.57, SE = 0.09, t = 6.33, p < .001$), followed by perceived anonymity ($\beta = .29, B = 0.48, SE = 0.11, t = 4.36, p < .001$). Both predictors made statistically significant unique contributions to the prediction of cyberbullying above and beyond the contribution of the other. Variance inflation factors ($VIF = 1.27$) confirmed the absence of problematic multicollinearity. These results provide full support for H4.

Table 5

Multiple Regression Analysis Predicting Cyberbullying (N = 300)

Predictor	B	SE B	β	t	p
Online Disinhibition	0.57	0.09	.43	6.33	<.001
Perceived Anonymity	0.48	0.11	.29	4.36	<.001
Model: $R^2 = .41$, Adjusted $R^2 = .40, F(2, 297) = 102.85, p < .001$					

Note. B = unstandardized regression coefficient; SE B = standard error; β = standardized regression coefficient. Dependent variable: Cyberbullying Inventory total score.

DISCUSSION

The present study investigated the roles of online disinhibition and perceived anonymity as predictors of cyberbullying perpetration among university students in Khyber Pakhtunkhwa, Pakistan. All four hypotheses were supported. Bivariate correlations confirmed significant positive associations among all three study variables, and multiple regression analysis demonstrated that online disinhibition and perceived anonymity together explained a substantial and theoretically meaningful 41% of the variance in cyberbullying perpetration. The findings extend the existing empirical literature to a previously understudied South Asian university population and provide culturally situated support for established theoretical frameworks.

The strong positive association between online disinhibition and cyberbullying ($r = .58$) constitutes the most robust finding of the study and is consistent with both theoretical predictions derived from Suler's (2004) Online Disinhibition Effect Theory and prior empirical evidence from other cultural contexts (Schacter et al., 2023; Udris, 2014; Wright, 2014). The regression coefficient ($\beta = .43$) further demonstrates that online disinhibition retains significant predictive power over cyberbullying even when perceived anonymity is statistically controlled, indicating that it is not merely an artifact of anonymity-mediated effects but an independent psychological determinant of harmful online behavior. Theoretically, this finding is consistent with Suler's account of toxic disinhibition as a multifactorial phenomenon driven by the convergence of anonymity, invisibility, asynchronicity, and the attenuation of normative authority—suggesting that online platforms as currently configured in KPK universities provide sufficient structural conditions for disinhibitory processes to operate.

The significant positive association between perceived anonymity and cyberbullying ($r = .49$, $\beta = .29$) is equally consistent with predictions derived from Barlett and Gentile's (2012) general theory of online aggression and the broader deindividuation literature (Festinger et al., 1952; Zimbardo, 1969). This finding suggests that university students in KPK who experience online interactions as anonymous are correspondingly more likely to engage in cyberbullying, presumably because reduced anticipated accountability lowers the psychological cost of aggressive behavior. This interpretation aligns with Barlett's (2015) longitudinal evidence that perceived anonymity temporally predicts increases in cyberbullying, and with the extensive experimental literature demonstrating that identified conditions produce substantially lower rates of online aggression than anonymous conditions (Lapidot-Lefler & Barak, 2012). Importantly, the unique contribution of perceived anonymity ($\beta = .29$) beyond online disinhibition in the regression model indicates that these two constructs, while correlated, capture distinct variance in cyberbullying—consistent with Suler's (2004) theoretical positioning of anonymity as a component mechanism of disinhibition rather than a simple proxy for it.

The significant positive correlation between perceived anonymity and online disinhibition ($r = .46$) provides empirical support for Suler's (2004) theoretical claim that anonymity is a constitutive mechanism of the disinhibition process. This finding suggests that students who perceive their online identity as concealed or unrecognizable are not merely more likely to engage in cyberbullying directly, but also more likely to experience a broader disinhibited orientation toward online communication that elevates cyberbullying risk through multiple pathways. The additive effect documented by Lapidot-Lefler and Barak (2012)—whereby anonymity and invisibility jointly produce greater toxic disinhibition than either alone—is theoretically consistent with this pattern. In the context of KPK, where the contrast between face-to-face social norms (characterized by collectivist accountability and familial surveillance) and the apparent freedom of anonymous online spaces may be particularly pronounced, the psychological amplification of disinhibitory effects by anonymity may be especially powerful.

The finding that 59% of variance in cyberbullying remained unexplained by the model underscores the multidetermined nature of cyberbullying perpetration and points toward additional psychological, social, and contextual variables warranting investigation. Candidates suggested by the broader literature include trait aggression (Kowalski et al., 2014), moral disengagement (Bandura, 1999), dispositional empathy (Udris, 2014), peer influence and network norms (Kelly et al., 2011), prior victimization history (Mishna et al., 2020), revenge motivation, family climate, and time spent online. Future research in this population should incorporate these variables to develop a more comprehensive predictive model.

Clinical and Policy Implications

The findings carry direct implications for university administrators, mental health professionals, and policymakers in Pakistan. First, the significant contribution of perceived anonymity to cyberbullying suggests that digital literacy programs emphasizing that online behavior is not truly anonymous—given institutional monitoring capabilities, digital forensics, and the traceability of online communications—may serve as an effective preventive strategy by correcting the misperception of absolute online anonymity. Second, the strong role of online disinhibition points toward the value of psychoeducational interventions that cultivate metacognitive awareness of the disinhibiting effects of online environments and build skills for self-regulation in digital contexts. Third, universities should establish clearly articulated and well-publicized anti-cyberbullying policies, including accessible reporting mechanisms, victim support services, and graduated sanctions for perpetration. Fourth, university counseling services should develop targeted interventions for both victims and perpetrators of cyberbullying, with particular attention to emotional regulation, empathy enhancement, and responsible digital communication. These recommendations are consistent with evidence-based practices advocated in the broader cyberbullying intervention literature (Mishna et al., 2020).

Limitations and Future Directions

The present study has several limitations that should be acknowledged. First, the cross-sectional design precludes causal inference; the direction of effects among online disinhibition, perceived anonymity, and cyberbullying cannot be established from correlational data. Future research employing longitudinal or daily diary designs would permit stronger inferences regarding temporal ordering and causal pathways. Second, the convenience sample—drawn from accessible students at universities in KPK—may not be representative of the broader KPK university student population, and findings may not generalize to students in other Pakistani provinces, rural areas, or those not enrolled in formal higher education. Third, reliance on self-report measurement introduces the possibility of response bias, including social desirability and underreporting of cyberbullying perpetration given the socially sensitive nature of the behavior. Future research might supplement self-report data with peer nominations or behavioral trace data where ethically feasible. Fourth, the study did not examine gender differences in the relationships among the study variables, despite growing evidence that online disinhibition and cyberbullying manifest differently across gender (Faucher et al., 2014; Greenfield et al., 2010). Moderation analyses examining gender as a boundary condition should be prioritized in future work. Fifth, the cultural generalizability of the psychometric instruments employed—all developed in Western or East Asian contexts—should be examined through formal adaptation and validation studies in the Pakistani context.

CONCLUSION

The present study provides empirical evidence that online disinhibition and perceived anonymity are significant positive predictors of cyberbullying perpetration among university students in Khyber Pakhtunkhwa, Pakistan. Together, these two psychological constructs accounted for 41% of the variance in cyberbullying behavior, demonstrating substantial predictive power and confirming the cross-cultural

applicability of Suler's (2004) Online Disinhibition Effect Theory and Barlett and Gentile's (2012) general theory of online aggression in a previously unstudied South Asian population. Online disinhibition emerged as the stronger predictor, followed by perceived anonymity, with both making independent contributions above and beyond each other. The findings underscore the urgency of developing evidence-based digital ethics education, anti-cyberbullying policies, and counseling interventions in Pakistani universities, and call for sustained investment in contextually appropriate cyberbullying prevention and response infrastructure. Future longitudinal and mixed-methods research is needed to clarify causal mechanisms and to identify additional psychological, social, and cultural factors that shape cyberbullying dynamics among Pakistani university students.

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
- Barlett, C. P. (2015). Anonymously hurting others online: The effect of anonymity on cyberbullying frequency. *Psychology of Popular Media Culture*, 4(2), 70–79. <https://doi.org/10.1037/ppm0000024>
- Barlett, C. P., & Gentile, D. A. (2012). Attacking others online: The formation of cyberbullying in late adolescence. *Psychology of Popular Media Culture*, 1(2), 123–135. <https://doi.org/10.1037/a0028113>
- Barlett, C. P., Gentile, D. A., & Chew, C. (2016). Predicting cyberbullying from anonymity. *Psychology of Popular Media Culture*, 5(2), 171–180. <https://doi.org/10.1037/ppm0000055>
- Beran, T., & Li, Q. (2007). The relationship between cyberbullying and school bullying. *Journal of Student Wellbeing*, 1(2), 15–33. <https://doi.org/10.21913/JSW.v1i2.172>
- Christopherson, K. M. (2007). The positive and negative implications of anonymity in Internet social interactions: "On the Internet, nobody knows you're a dog." *Computers in Human Behavior*, 23(6), 3038–3056. <https://doi.org/10.1016/j.chb.2006.09.001>
- Faucher, C., Jackson, M., & Cassidy, W. (2014). Cyberbullying among university students: Gendered experiences, impacts, and perspectives. *Education Research International*, 2014, Article 698545. <https://doi.org/10.1155/2014/698545>
- Festinger, L., Pepitone, A., & Newcomb, T. (1952). Some consequences of deindividuation in a group. *Journal of Abnormal and Social Psychology*, 47(2), 382–389. <https://doi.org/10.1037/h0057906>
- Gorzig, A., & Machackova, H. (2015). Cyberbullying from a sociocultural perspective: Central and Eastern European perspectives. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 9(4), Article 1. <https://doi.org/10.5817/CP2015-4-1>
- Hinduja, S., & Patchin, J. W. (2008). Cyberbullying: An exploratory analysis of factors related to offending and victimization. *Deviant Behavior*, 29(2), 129–156. <https://doi.org/10.1080/01639620701457816>
- Kowalski, R. M., Giumetti, G. W., Schroeder, A. N., & Lattanner, M. R. (2014). Bullying in the digital age: A critical review and meta-analysis of cyberbullying research among youth. *Psychological Bulletin*, 140(4), 1073–1137. <https://doi.org/10.1037/a0035618>

- Lapidot-Lefler, N., & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, 28(2), 434–443. <https://doi.org/10.1016/j.chb.2011.10.014>
- Mishna, F., Cook, C., Saini, M., Wu, M. J., & MacFadden, R. (2020). Interventions to prevent and reduce cyber abuse of youth: A systematic review. *Research on Social Work Practice*, 21(1), 5–14. <https://doi.org/10.1177/1049731509351988>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Pakistan Telecommunication Authority. (2023). Annual report 2022–23. <https://www.pta.gov.pk>
- Schacter, H. L., Varma, A., & Juvonen, J. (2023). Cyberbullying and online disinhibition: Examining links with adolescent social anxiety and hostile attribution bias. *Journal of Early Adolescence*, 43(3), 327–351. <https://doi.org/10.1177/02724316221078817>
- Schenk, A. M., & Fremouw, W. J. (2012). Prevalence, psychological impact, and coping of cyberbully victims among college students. *Journal of School Violence*, 11(1), 21–37. <https://doi.org/10.1080/15388220.2011.630310>
- Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry*, 49(4), 376–385. <https://doi.org/10.1111/j.1469-7610.2007.01846.x>
- Sticca, F., & Perren, S. (2013). Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying. *Journal of Youth and Adolescence*, 42(5), 739–750. <https://doi.org/10.1007/s10964-012-9867-3>
- Suler, J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*, 7(3), 321–326. <https://doi.org/10.1089/1094931041291295>
- Tanrikulu, I., & Erdur-Baker, O. (2021). Psychometric properties of a Cyberbullying Inventory for university students. *British Journal of Guidance & Counselling*, 49(1), 68–80. <https://doi.org/10.1080/03069885.2020.1775174>
- Udris, R. (2014). Cyberbullying among high school students in Japan: Development and validation of the Online Disinhibition Scale. *Computers in Human Behavior*, 41, 253–261. <https://doi.org/10.1016/j.chb.2014.09.036>
- Willard, N. E. (2007). *Cyberbullying and cyberthreats: Responding to the challenge of online social aggression, threats, and distress*. Research Press.
- Wright, M. F. (2014). Predictors of anonymous cyber aggression: The role of adolescents' beliefs about anonymity, aggression, and the permanency of digital content. *Cyberpsychology, Behavior, and Social Networking*, 17(7), 431–438. <https://doi.org/10.1089/cyber.2013.0457>
- Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. *Nebraska Symposium on Motivation*, 17, 237–307.