

Toxic Echoes: The Cyberpsychology of Cancel Culture within Digital Student Communities

Juni Hartiwi^{*}, Sri sutiwi²

¹STKIP PGRI Bandar Lampung, Indonesia

²ITBA DCC Bandar Lampung, Indonesia

* Email: juni_marvel@yahoo.com

Submitted: 10/02/2026	<p style="text-align: center;">Abstract</p> <p>Cancel culture has become a dominant form of social regulation within digital student communities, profoundly impacting the psychological well-being and academic trajectories of university students. This systematic narrative literature review synthesizes current research on the cyberpsychological dimensions of cancel culture among young adults. Drawing on frameworks such as the online disinhibition effect, space transition theory, and moral grandstanding, this review identifies the structural and psychological drivers of online mob behavior. It examines the severe mental health consequences for targets, including toxic shame, depression, and academic sabotage, exacerbated by the digital permanence of public shaming. Furthermore, the review explores the collective anxiety and self-censorship experienced by bystanders, as well as the moral grandstanding fatigue suffered by active participants. Finally, the study highlights critical gaps in the existing literature, such as the lack of longitudinal and intersectional analyses, and provides recommendations for educational institutions to cultivate digital empathy, implement restorative justice frameworks, and safeguard academic freedom.</p> <p>Keywords: Cancel culture, cyberpsychology, digital student communities, online disinhibition, cyberbullying, student mental health, academic sabotage, virtue signaling.</p>
Accepted: 01/03/2026	
Published: 31/03/2026	

© The Author(s). 2026 Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0>

INTRODUCTION

The accelerating digitalization of higher education has fundamentally restructured the social landscape that students inhabit. University life is no longer confined to physical campuses; it has extended into a vast ecosystem of online platforms, Reddit communities, TikTok, Discord servers, and Instagram group chats, that operate continuously and across geographic boundaries. As these digital student communities have grown in scale and intimacy, they have also become powerful arenas of social judgment. Within this context, "cancel culture", broadly understood as the collective act of publicly shaming, deplatforming, or socially ostracizing an individual perceived to have violated group norms, has emerged as one of the most dominant forms of peer regulation in student life today. The prevalence and intensity of this phenomenon

demand rigorous academic inquiry, particularly from the lens of cyberpsychology, which examines how humans behave, think, and feel when interacting through digital technology (Suler, 2004). Understanding the roots and consequences of cancel culture within student communities is not merely an academic exercise; it is a matter of urgent psychological and educational concern.

The relevance of this topic is underscored by a well-documented global rise in mental health challenges among university students. Research consistently shows that social media environments, the same platforms on which cancel culture proliferates, are strongly associated with elevated symptoms of depression, anxiety, and psychological stress (Shannon et al., 2022; Karim et al., 2020). A large-scale systematic review and meta-analysis involving thousands of participants found statistically significant correlations between problematic social media use and depression, anxiety, and stress, with effects consistent across age groups and genders (Shannon et al., 2022). Cancel culture amplifies these risks by introducing targeted, community-wide condemnation into already psychologically volatile digital spaces. When students, who are in a critical period of identity formation and social development, become targets of public shaming campaigns within their peer networks, the psychological consequences can extend far beyond the digital domain, impairing academic functioning and real-world social relationships (Steinsbekk et al., 2023).

Cancel culture occupies a contested ideological position that makes it particularly difficult to address within educational institutions. It originated as, and continues to function as, a mechanism of social accountability, allowing marginalized groups to exercise collective power against those who cause harm (Ng, 2022). Empirical research published in *Frontiers in Psychology* confirms that for groups experiencing systemic harm, cancel culture can generate feelings of collective validation and moral solidarity, reinforcing its appeal as a justice mechanism (Traversa et al., 2023). However, the same study acknowledges that cancel culture is routinely criticized when it escapes proportionality and targets individuals based on incomplete information, in-group conformity pressure, or moral panic rather than genuine accountability (Traversa et al., 2023). In insular digital student communities, where anonymity lowers inhibition, echo chambers amplify outrage, and social hierarchies are acutely felt, the mechanism frequently mutates from justice-seeking into a form of coordinated cyberbullying, producing material and mental health losses for targets (Laili et al., 2024). This duality, between social corrective and psychological weapon, is precisely what makes cancel culture such a critical subject of investigation in the current academic climate.

The intersection of digital anonymity, adolescent developmental psychology, and online platform architecture creates what researchers have called a "disinhibition effect," in which individuals engage in behaviors online that they would suppress in face-to-face interaction (Suler, 2004). This disinhibition has been identified as a core driver of cyberbullying and online harassment, phenomena that share structural and motivational similarities with cancel culture (Laili et al., 2024; Julianti, 2023). Research examining cancel culture through the lens of Space Transition Theory, which posits that individuals carry suppressed behaviors from physical to virtual spaces, found that emotional triggers such as value conflicts, disappointment, and poor self-regulation are primary antecedents of cancel culture actions on social media (Julianti, 2023). For student populations, who are navigating identity negotiation, social belonging, and academic pressure simultaneously, these dynamics create a uniquely toxic psychosocial environment. Bystanders to cancel culture events are equally affected; exposure to coordinated

online aggression within one's community generates vicarious anxiety and a chilling effect on authentic self-expression (Picarella, 2024).

The growing convergence of cancel culture, misinformation, and platform society dynamics further elevates the urgency of this research. Picarella (2024) argues that the transition from a "network society" to a "platform society" has structurally enabled cancel culture by concentrating social interaction on algorithmic platforms that reward outrage and virality over nuance and deliberation. Misinformation and the instantaneous nature of social media are identified as primary accelerants: false or decontextualized information circulates rapidly within student communities, triggering cancel events before verification is possible (Laili et al., 2024). Concurrently, survey experiments examining the political psychology of cancel culture reveal that individuals' support for or opposition to canceling is shaped not merely by values, but by in-group identity and partisan loyalty, suggesting that cancel culture in student communities may function as a social sorting mechanism that reinforces ideological homogeneity rather than promoting genuine accountability (Diether et al., 2024). Given that higher education institutions are specifically mandated to cultivate critical thinking, open discourse, and psychological safety, the unchecked proliferation of cancel culture in digital student communities represents a direct threat to the foundational mission of the university. This research therefore contributes to the emerging scholarly conversation on digital citizenship, campus mental health, and the ethical governance of online student communities, areas that remain understudied despite their growing practical and theoretical significance.

METHOD

This study employs a systematic narrative literature review as its primary research design, structured to synthesize existing empirical and theoretical scholarship on the cyberpsychological dimensions of cancel culture within digital student communities. A systematic narrative review combines the transparency and reproducibility of systematic methodology with the interpretive flexibility necessary to integrate diverse theoretical frameworks from psychology, sociology, and media studies (Page et al., 2021). The review was conducted in accordance with the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, which provide a standardized 27-item checklist designed to ensure complete, transparent, and accurate reporting of the review process, including study identification, screening, eligibility assessment, and inclusion (Page et al., 2021). Literature searches were conducted primarily through Google Scholar, supplemented by databases including PsycINFO, Scopus, and ERIC, to maximize coverage of peer-reviewed psychological, sociological, and educational research. The following keyword combinations were used systematically across all databases: "cancel culture" AND "student mental health"; "cyberbullying" AND "university"; "online public shaming" AND "psychology"; "digital ostracism" AND "adolescents"; "social media" AND "peer aggression" AND "higher education"; and "cyberpsychology" AND "online disinhibition" AND "young adults". Boolean operators (AND, OR) and truncation techniques were applied to maximize search sensitivity while maintaining specificity to the study's core constructs.

The selection of sources was governed by clearly defined inclusion and exclusion criteria to ensure relevance, quality, and population specificity. For inclusion, studies were required to be: (1) peer-reviewed journal articles, book chapters from academic publishers, or empirically grounded sociological or psychological analyses; (2) published within the last ten years (2015–2025), with exceptions made for seminal theoretical works such as Suler's (2004) foundational work on the online disinhibition effect; (3) focused on high school or university-aged populations (approximately 15–25 years), or on general adolescent and young adult digital behavior directly applicable to student communities; and (4) addressing at least one of the

following core constructs: cancel culture, online public shaming, cyberbullying, digital peer aggression, social ostracism, or the psychological effects thereof. Studies were excluded if they focused exclusively on celebrity cancel culture without psychological analysis, corporate public relations crises, political cancel culture with no bearing on student populations, or non-digital forms of bullying that lacked relevance to online peer dynamics. Following the screening and selection process, the retained literature was subjected to reflexive thematic analysis as the primary data analysis method. Thematic analysis, as developed by Braun and Clarke (2006), involves a systematic six-phase process of familiarization with data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the final report, allowing for rich, interpretive synthesis across qualitative and quantitative sources (Braun & Clarke, 2023). The identified literature was coded and categorized into three primary psychological theme clusters: (1) drivers and motivations of cancel culture behavior among students; (2) psychological impacts on targets, perpetrators, and bystanders; and (3) contextual amplifiers, including digital anonymity, echo chambers, and platform architecture. This thematic structure provides the organizational backbone of the review's findings and discussion sections.

RESULT AND DISCUSSION

Theme 1: The Cyberpsychological Drivers of the "Mob"

Online Disinhibition Effect

One of the most foundational psychological mechanisms driving cancel culture behavior within digital student communities is the online disinhibition effect, a phenomenon whereby the perceived anonymity, invisibility, and asynchronicity of online communication systematically reduce the inhibitions that regulate aggressive behavior in face-to-face contexts (Suler, 2004). Suler (2004) identifies two forms of this effect: benign disinhibition, in which users share vulnerabilities and emotions more freely, and toxic disinhibition, in which individuals engage in hostility, harassment, and derogatory language that they would suppress in offline social settings. Cancel culture in student communities operates predominantly through the latter mechanism: the screen functions as a psychological buffer that depersonalizes the target, reduces empathy toward their suffering, and emboldens participants who would otherwise fear social repercussions (Van Grunsven & Marin, 2024). Research examining cancel culture through this lens confirms that digital platforms create "affordances hostile to healthy, risky interpersonal identity-exploration," fundamentally distorting the social-developmental function of peer interaction for adolescents and young adults (Van Grunsven & Marin, 2024).

The structural features of social media platforms, algorithmic amplification, public comment threads, and viral sharing mechanics, transform individual acts of disinhibition into coordinated mob behavior at extraordinary speed. Julianti (2023) found, using Space Transition Theory, that cancel culture actors on Twitter are motivated by a convergence of emotional triggers including disappointment, value collisions, and the perceived ease of mass mobilization afforded by digital anonymity, factors that override self-regulatory impulses that would otherwise moderate behavior. Critically, Jochan and Banerjee (2021) situate this within a broader framework of digital vigilantism, arguing that online platforms have become instruments of disciplinary surveillance in which social shaming is deployed as a mechanism of community policing, often without due process, proportionality, or concern for the psychological consequences for the target. This architecture of mob formation is particularly potent among student populations,

whose developmental need for peer validation makes them both likely participants in and highly vulnerable targets of digital pile-ons (Chonge et al., 2025).

The disinhibition effect is further compounded by what researchers term the deindividuation of online crowd behavior, in which individuals submerge their personal identity into the collective identity of the canceling mob, reducing personal accountability and increasing conformity to the group's punitive impulses (Suler, 2004). Muir et al. (2023) identify moral disengagement, a psychological process in which individuals rationalize harmful behavior by framing it as morally justified punishment, as a key predictor of online shaming behavior, explaining why students who would never confront a peer aggressively in a seminar room readily participate in large-scale public condemnation online. Mueller (2021) further documents that the most significant behavioral predictor of overall cancel culture involvement is the immediate demand for a public apology, a finding that reveals the performative and punitive, rather than genuinely restorative, character of cancel culture as it operates in peer communities. Together, these findings establish that the "mob" is not a random social event but the product of specific, identifiable cyberpsychological conditions that digital student environments reliably reproduce.

Virtue Signaling and Identity Formation

The participation of young adults in cancel culture is not solely driven by aggression or disinhibition; it is also deeply rooted in the developmental psychology of identity formation and the need for social belonging. Erikson's framework of psychosocial development identifies late adolescence and early adulthood as the critical period during which individuals resolve the tension between identity and role confusion, a process that is increasingly mediated by digital social environments for contemporary students (Van Grunsven & Marin, 2024). In this context, publicly aligning oneself with the group's moral consensus through acts of cancel culture serves as a powerful identity-signaling mechanism: by joining the condemnation of a perceived transgressor, the student communicates shared values to the in-group and secures their position within the community's social hierarchy (Diether et al., 2024). Mueller (2021) found that involvement in cancel culture was significantly associated with political liberalism and identification with feminine gender traits, suggesting that the behavior functions as a form of ideological self-expression and social positioning rather than purely punitive action.

This dynamic is closely related to the concept of moral grandstanding, the use of public moral discourse primarily to signal one's own virtue or goodness to an audience, rather than to address genuine ethical wrongdoing (Muir et al., 2023). Muir et al. (2023) empirically demonstrate that moral grandstanding is a statistically significant predictor of online shaming intentions, establishing a direct link between the psychological need for social status and participation in cancel campaigns. The perceived social rewards of publicly condemning a peer, increased group approval, affirmation of one's moral identity, and enhanced social capital, create a powerful incentive structure that makes cancel culture self-reinforcing within close-knit student communities such as university course groups, club Discord servers, or campus social media pages. Ng (2022) argues that this dynamic produces an environment where social norms are enforced not through deliberative ethical reasoning, but through competitive displays of moral indignation, fundamentally distorting the epistemic culture that universities are designed to foster.

The long-term consequence of virtue signaling as a driver of cancel culture is the emergence of a performative moral economy within digital student communities, a social system in which

the currency of group membership is demonstrable moral outrage, and in which nuance, context, and proportionality are sacrificed in favor of legibility as a "good person" to one's peers (Alkhateeb & Bouherar, 2024). Alkhateeb and Bouherar (2024) document how this dynamic manifests within higher education, finding that academics and students alike adopt a "culture of safetyism", the preemptive restriction of discourse, expression, and intellectual risk-taking in order to avoid becoming targets of moral condemnation. Chong et al. (2025) observe that this psychosocial pressure is particularly acute for students because their sense of belonging, academic identity, and social survival are all located within the same community, meaning that social exclusion carries consequences that extend across every dimension of their university experience. The intersection of developmental identity needs and digital moral economies therefore creates a structural condition in which virtue signaling is both psychologically motivated and socially rewarded, sustaining cancel culture as a persistent feature of digital student life.

Echo Chambers and Polarization

The algorithmic architecture of the social media platforms most popular among students, TikTok, Instagram, Reddit, and Twitter/X, is engineered to maximize engagement by prioritizing content that confirms users' existing beliefs and amplifies emotionally resonant material, producing the well-documented phenomenon of the echo chamber (Picarella, 2024). Within these algorithmically curated environments, student communities develop increasingly homogeneous ideological profiles, in which shared moral positions are not tested against diverse perspectives but are instead continuously reinforced, radicalized, and rendered non-negotiable (Picarella, 2024). Diether et al. (2024) demonstrate empirically that support for or opposition to cancel culture events is driven less by abstract values than by in-group identity and partisan loyalty, meaning that within a polarized student echo chamber, the primary question is not "was this action genuinely harmful?" but rather "is this person one of us?" Minor infractions, stripped of context and filtered through an outrage-amplification algorithm, are thus reframed as catastrophic moral violations that demand immediate communal response.

The echo chamber effect is further intensified by the structural characteristics of closed student digital communities, university course group chats, faculty-specific Discord servers, and institution-tagged subreddits, which combine the intimacy of a known peer group with the disinhibiting anonymity of online interaction. Traversa et al. (2023) find that within these contexts, cancel culture produces strong feelings of collective validation for in-group members, reinforcing the perceived righteousness of the mob's actions and generating collective empowerment that sustains further engagement. This creates a feedback loop in which the initial act of cancellation generates in-group solidarity, which in turn lowers the threshold for future cancellations and increases the severity of punitive responses. The target, viewed through the distorting lens of echo chamber dynamics, is perceived not merely as an individual who made a mistake, but as a symbol of the ideological threat that the group has collectively defined itself against (Van Grunsven & Marin, 2024).

Critically, the polarization produced by echo chambers does not only affect the target of a cancel event, it degrades the epistemic environment of the entire student community. Edling (2026) examines how dogma within higher education settings risks "hindering academic freedom and stifling research discussions," arguing that the enforced ideological conformity characteristic of campus cancel culture creates an environment in which students and academics are reluctant to test unconventional ideas, challenge consensus positions, or engage in the productive disagreement that is the hallmark of genuine intellectual inquiry. Alkhateeb and Bouherar (2024)

corroborate this finding across Arab university contexts, demonstrating that cancel culture events generate a pervasive caution among academics and students that constricts the boundaries of legitimate discourse. The cumulative effect of echo chambers and polarization in digital student communities is therefore not merely the cancellation of individual students, but the systematic erosion of the university's core epistemic function, the free, evidence-based pursuit of knowledge, replaced by a digital social order in which ideological conformity is the primary criterion of community membership.

Table 1: Cyberpsychological Drivers of Cancel Culture in Digital Student Communities

Driver	Core Mechanism	Key Theoretical Basis	Primary Effect on Students
Online Disinhibition Effect	Anonymity & asynchronicity reduce empathy and inhibition	Suler (2004); Jochan & Banerjee (2021)	Increased aggression, mob formation, digital vigilantism
Virtue Signaling	Public moral alignment to secure in-group status	Mueller (2021); Muir et al. (2023)	Performative condemnation, competitive moral outrage
Moral Grandstanding	Shaming used to signal personal virtue	Muir et al. (2023); Ng (2022)	Punitive rather than restorative social policing
Identity Formation Need	Belonging secured through ideological conformity	Van Grunsven & Marin (2024); Diether et al. (2024)	Cancel acts as social membership signal
Echo Chamber Dynamics	Algorithms reinforce ideological homogeneity	Picarella (2024); Traversa et al. (2023)	Minor infractions reframed as catastrophic violations
Collective Validation	In-group solidarity reinforced by shared condemnation	Traversa et al. (2023)	Self-reinforcing cycle of cancellation

Theme 2: Psychological Impact on the "Canceled" (The Targets)

Mental Health Crisis

The psychological harm inflicted on students who become targets of cancel culture within digital communities is severe, multidimensional, and clinically significant. At the most acute level, research on cyber-victimization among university students consistently documents elevated rates of depression, generalized anxiety disorder, and suicidal ideation as direct consequences of peer-based online ostracism (Hossain et al., 2023; Karim et al., 2020). A study of 1,509 Chinese college students found that cyberbullying victimization was significantly associated with suicidal ideation, mediated by core self-evaluation deficits and depressive symptoms, establishing a clear empirical pathway from digital peer aggression to the most severe mental health outcomes (Li et al., 2022). Cancel culture intensifies these effects by combining the harm of cyberbullying with the additional psychological burden of perceived community-wide moral rejection, the target is not merely harassed, but condemned as fundamentally unworthy of social

belonging by the very community from which they derive their identity, academic purpose, and daily social support (Chonge et al., 2025).

Ungerer (2021) identifies the concept of toxic shame as the core psychological mechanism through which cybershaming, the digital analogue of cancel culture, produces its most lasting damage: unlike guilt, which attaches to a specific behavior, toxic shame attaches to the self, generating the belief that one is fundamentally defective, unworthy, and deserving of rejection. This distinction is critical for understanding why cancel culture targeting has clinical consequences that outlast the duration of the online event itself: the internalized verdict of the community, delivered publicly and at scale, becomes a persistent component of the target's self-concept (Ungerer, 2021). Jochan and Banerjee (2021) document through survey data that students who experience digital shaming report significant alterations in their online and offline behavior, including social withdrawal, hypervigilance in digital communication, and reluctance to participate in campus life, behavioral signatures consistent with post-traumatic stress responses. Chonge et al. (2025) further establish that cancel culture functions as a profound threat to students' sense of belonging and identity, with ostracized students reporting disrupted affiliations with their academic institutions, peer groups, and sense of future purpose.

The mental health crisis produced by cancel culture targeting is also mediated by the unique power asymmetry inherent in collective condemnation: a single student faces the coordinated judgment of an entire networked community, with no formal process for defense, contextualization, or appeal (Lokhande & Natu, 2022). Lokhande and Natu (2022) document through interviews with university faculty and associate professors that the experience of social exclusion from digital spaces produces in student targets a profound sense of helplessness and loss of control, psychological states strongly associated with the development of clinical depression and anxiety. Shannon et al. (2022) establish through systematic review and meta-analysis that problematic engagement with social media environments of the type that produce cancel events is independently predictive of clinically elevated mental health symptom scores across depression, anxiety, and stress subscales. The convergence of toxic shame, community-level rejection, and psychological helplessness creates a constellation of mental health risks that, in the most severe cases, places canceled students at elevated risk for self-harm and suicidal ideation, a finding that demands urgent attention from student welfare services and institutional policy-makers alike.

Academic Sabotage

The psychological distress produced by cancel culture targeting does not remain confined to the emotional domain; it translates directly into measurable impairment of academic functioning, creating what this review terms academic sabotage, the systematic disruption of a student's capacity to learn, perform, and persist in their educational trajectory as a result of social cancellation. Cyberbullying research consistently documents that victims experience significant declines in academic performance, increased absenteeism, and higher dropout rates compared to non-victimized peers, with cyber-victimized university students in Bangladesh showing significantly greater likelihood of sleep difficulties, reduced social media engagement, and elevated suicidal ideation that collectively undermine academic participation (Hossain et al., 2023). Cancel culture extends these harms by making the academic environment itself, seminars, group projects, campus events, feel socially hostile, as the target must navigate public condemnation by the very peers with whom they share their educational experience (Chonge et al., 2025).

The relationship between social ostracism and academic decline is well-supported in the psychological literature on belonging and motivation. Students who perceive that they have been expelled from the social community of their institution experience a collapse of academic motivation, as belonging to the institution's social fabric is a precondition for the investment of effort, participation, and identity in academic goals (Chonge et al., 2025). Van Grunsven and Marin (2024) note that Social Media Platform-afforded cancel culture specifically disrupts "the social relational domains essential for teenagers' development," and while their analysis centers on identity formation, the disruption of these relational domains has a direct corollary in academic engagement: students cannot learn effectively in an environment where their social presence is experienced as threatening. Alkhateeb and Bouherar (2024) find that in higher education contexts, the academic consequences of cancel culture extend to faculty as well as students, with canceled academics reporting avoidance of certain research topics, curtailment of pedagogical risk-taking, and reduced publication of findings on sensitive subjects, a pattern of academic self-sabotage that mirrors what students experience at the peer level.

The full extent of academic sabotage caused by cancel culture remains underquantified in the existing literature, partly because students who withdraw from courses or disengage from academic communities as a result of social cancellation are rarely identified as such in institutional records (Lokhande & Natu, 2022). Steinsbekk et al. (2023) provide longitudinal data showing that social media-associated anxiety and depression symptoms, tracked across ages 10–16, accumulate over time rather than resolving spontaneously, a finding with direct implications for the academic trajectories of university students who experience cancel events at a formative stage of their educational careers. The practical consequences are significant: a student who withdraws from a course following a cancel event, or who drops out of university due to the social and psychological consequences of public shaming, faces long-term career disadvantage that the existing literature does not yet systematically track. This gap, the absence of longitudinal data on the career and educational outcomes of cancel culture targets, represents one of the most urgent methodological deficits in the field, discussed further in Theme 4.

Digital Permanence

A feature of cancel culture in digital environments that has no analogue in offline peer conflict is the permanent, searchable, and shareable nature of the record it creates. Unlike a rumor that fades from memory or a public confrontation that loses witnesses over time, a cancel campaign conducted across social media platforms produces a documented archive of accusation, condemnation, and public judgment that persists indefinitely and can be surfaced by any future employer, colleague, academic institution, or personal contact through a routine online search (Ungerer, 2021; Jochan & Banerjee, 2021). Ungerer (2021) describes cybershaming as producing "damage to reputation and job loss" as direct downstream consequences of online sanctioning, establishing that the harm does not terminate when the immediate social furor subsides but continues to accumulate across a student's post-university life. This temporal dimension of digital harm is psychologically distinct from other forms of peer aggression and constitutes, in its own right, a form of extended punishment that is wholly disproportionate to any realistic assessment of the initial transgression.

The psychological impact of digital permanence is compounded by the context collapse that social media archives create: a message, post, or behavior that was contextually appropriate within its original social frame is stripped of that context when surfaced in a new search or shared by a new audience, making it perpetually vulnerable to reinterpretation and re-condemnation

(Picarella, 2024). Jochan and Banerjee (2021) describe this as a form of digital surveillance that extends the reach of disciplinary shame beyond any single community or moment in time, creating what they term a digital record of moral unworthiness that follows the individual across contexts. For student targets, who are at the beginning of their professional and social lives, this permanence is particularly devastating: the cancel event occurs at a moment of maximum identity vulnerability and minimum institutional power, and its digital record may define how they are perceived by future employers, professional communities, and social networks for decades (Lokhande & Natu, 2022).

The psychological trauma of digital permanence is also self-perpetuating: canceled students who are aware that their condemnation is permanently documented report hypervigilance in all online communication, avoidance of digital public spaces, and a pervasive sense of exposure and vulnerability that mirrors the symptom profile of post-traumatic stress disorder (Jochan & Banerjee, 2021). Ungerer (2021) introduces the concept of shame resilience as a potential protective factor, noting that individuals with high shame resilience, the capacity to recognize shame, process it without self-destructive internalization, and reconnect with community, are better equipped to withstand the psychological consequences of cybershaming. However, the very characteristics of digital student communities, their insularity, the developmental vulnerability of their members, and the absence of restorative institutional mechanisms, systematically undermine shame resilience, leaving targets without the social scaffolding necessary for recovery (Chonge et al., 2025). The permanence of digital condemnation thus represents not merely a reputational harm, but a chronic psychological stressor that demands specific therapeutic and policy responses from universities.

Table 2: Psychological Impact on Cancel Culture Targets, Summary of Evidence

Impact Domain	Documented Effects	Key Sources
Anxiety & Depression	Clinically elevated GAD-7 and PHQ-9 scores among cyber-victimized students	Hossain et al. (2023); Shannon et al. (2022)
Suicidal Ideation	Significant association between cyberbullying victimization and SI, mediated by depression	Li et al. (2022); Karim et al. (2020)
Toxic Shame	Shame attaches to self-concept, not behavior; produces persistent identity damage	Ungerer (2021)
Loss of Belonging	Disrupted affiliation with peers, institution, and sense of future purpose	Chonge et al. (2025)
Academic Decline	Absenteeism, withdrawal, reduced engagement, and performance impairment	Steinsbekk et al. (2023)
Digital Permanence	Permanent searchable record creates chronic reputational and psychological harm	Jochan & Banerjee (2021); Ungerer (2021)
Behavioral Alteration	Social withdrawal, hypervigilance, avoidance of campus life	Jochan & Banerjee (2021)

Theme 3: The Toll on Bystanders and the "Cancelers"

Collective Anxiety and Self-Censorship

The psychological consequences of cancel culture within digital student communities are not limited to those directly targeted; they propagate outward through the community in the form

of collective anxiety, a pervasive, community-wide apprehension generated by the observation of cancellation events that systematically suppresses authentic expression, intellectual risk-taking, and genuine participation in academic life. Bystanders who witness cancel campaigns within their social network rapidly internalize the implicit lesson that the community's social boundaries are enforced through swift, public, and disproportionate punishment, and adjust their own behavior accordingly, engaging in preemptive self-censorship to minimize their risk of becoming the next target (Edling, 2026). Edling (2026) documents this dynamic specifically within higher education, arguing that the threat of cancel culture produces a form of dogmatic conformity in academic settings, in which students and academics alike suppress engagement with contested ideas, heterodox perspectives, and sensitive research questions in order to avoid triggering community condemnation. The result is a chilling effect on the free intellectual inquiry that is the fundamental purpose of the university.

Alkhateeb and Bouherar (2024) provide empirical evidence for this dynamic in their study of cancel culture in Arab universities, finding that academics adopt a "culture of safetyism", the systematic avoidance of topics, approaches, and expression that might provoke community censure, as a rational response to the perceived threat of cancellation. While their focus is on faculty, the same logic applies with equal or greater force to students, whose institutional power is even more limited and whose social survival is entirely dependent on the goodwill of the peer community. Chong et al. (2025) establish that the psychosocial stress of cancel culture permeates the entire student social ecology, not merely the experiences of direct targets, creating an ambient climate of threat that impairs resilience and coping across the student body. This collective anxiety is psychologically costly not only because it suppresses expression, but because the continuous effort of self-monitoring and self-censorship is cognitively and emotionally exhausting, diverting psychological resources from learning, creativity, and genuine social connection.

The self-censorship produced by cancel culture anxiety has documented consequences for the quality of academic discourse and the development of critical thinking skills. Edling (2026) draws on John Dewey's framework to argue that genuine moral and intellectual development requires the freedom to make and recover from mistakes in a context of psychological safety, precisely the condition that cancel culture destroys. Van Grunsven and Marin (2024) propose that digital literacy education focused specifically on the "socioethical challenges of growing up online" is essential for mitigating the chilling effects of cancel culture in student communities, acknowledging that the problem cannot be addressed by individual behavioral change alone but requires institutional intervention in how students are taught to navigate digital social environments. The bystander experience of cancel culture is therefore not a passive one: it is a continuous, psychologically costly process of social navigation in an environment where the penalties for miscalculation are publicly applied and permanently recorded, fundamentally altering the conditions of possibility for authentic intellectual and social engagement among students.

Moral Grandstanding Fatigue

A less-studied but psychologically significant consequence of cancel culture is the psychological toll experienced by the cancelers themselves, the students who actively participate in condemnation campaigns and who, over time, experience a particular form of cognitive and emotional exhaustion that this review terms moral grandstanding fatigue. Mueller (2021) identifies demanding a public apology as the most powerful predictor of cancel culture

involvement, a finding that positions cancel culture participants primarily as moral enforcers whose engagement is sustained by the expectation of visible contrition from the target. However, when targets fail to perform the expected apology or when the social reward of condemnation proves short-lived, participants are left with the psychological residue of their own aggression, a dissonance that can generate anxiety, regret, and a growing fear of reputational vulnerability. Muir et al. (2023) find that moral grandstanding and moral disengagement are co-present predictors of shaming behavior, suggesting that cancel culture participants are simultaneously seeking social validation and rationalizing away the harm they cause, a psychologically unstable combination that cannot be sustained without cost.

The fear of becoming the next target is a specific and well-documented source of anxiety for active cancel culture participants (Chonge et al., 2025). Students who have participated in cancel events are acutely aware that their own digital histories, social positions, and expressed opinions are equally subject to retrospective condemnation, and that the community they currently control can turn against them with the same speed and mercilessness that they directed at others. Ungerer (2021) notes that cybershaming communities operate according to shifting and internally inconsistent norms, making it impossible for any participant to be certain of long-term immunity from condemnation. This epistemological insecurity generates a form of hypervigilant self-monitoring among active participants that Chonge et al. (2025) describe as psychosocially debilitating, the individual must continuously surveil both their own behavior and the shifting moral consensus of the community in order to maintain their position, a cognitive burden that accumulates into genuine psychological exhaustion.

Table 3: Psychological Consequences Across Campus Community Roles

Role	Primary Psychological Effects	Behavioral Manifestation	Key Sources
Target (Canceled)	Toxic shame, depression, anxiety, suicidal ideation, PTSD symptoms	Social withdrawal, academic decline, digital avoidance	Ungerer (2021); Chonge et al. (2025); Li et al. (2022)
Bystander	Collective anxiety, self-censorship, suppressed expression	Avoidance of controversial topics, reduced academic participation	Edling (2026); Alkhateeb & Bouherar (2024)
Canceler (Perpetrator)	Moral grandstanding fatigue, paranoia, fear of retribution	Hypervigilant self-monitoring, interpersonal distrust	Mueller (2021); Muir et al. (2023); Chonge et al. (2025)

The long-term psychological profile of habitual cancel culture participants has received minimal systematic investigation, representing a significant gap in the literature. However, existing evidence suggests that moral grandstanding fatigue manifests as a cluster of symptoms including emotional exhaustion, interpersonal distrust, cynicism about online social environments, and a paradoxical sense of social isolation, ironic given that cancel culture behavior is ostensibly performed in the service of community cohesion (Muir et al., 2023). Traversa et al. (2023) find that for marginalized group members, participation in cancel culture generates collective empowerment and collective anger/contempt, emotional states that are psychologically activating in the short term but likely to produce depletion and burnout with repeated exposure, particularly when the systemic conditions that motivated the initial grievance remain unaddressed. The result is a digital student community in which cancelers, bystanders,

and targets are all, in different ways, experiencing chronic psychological distress, a finding that underscores the argument that cancel culture is not a tool of social health but a systemic psychological stressor affecting every member of the communities in which it operates.

Theme 4: Synthesis of Gaps in the Literature

The systematic review of existing scholarship reveals a body of literature that has made important foundational contributions to understanding cancel culture as a cyberpsychological and sociological phenomenon, but which contains several critical methodological and conceptual gaps that constrain both theoretical development and practical intervention design. The most significant of these is the near-complete absence of longitudinal research tracking the long-term psychological and career consequences of college-era cancellation. The studies reviewed, including Mueller (2021), Muir et al. (2023), and Chonge et al. (2025), are predominantly cross-sectional, capturing the psychological state of participants at a single point in time and providing no data on whether the mental health consequences of being canceled resolve, persist, or intensify over months and years. Steinsbekk et al.'s (2023) four-wave longitudinal cohort study of social media-associated anxiety and depression in younger adolescents represents an important methodological model, but no equivalent study has been conducted specifically for university-age cancel culture experiences.

A second major gap is the absence of intersectional analysis examining which student populations bear the greatest burden of cancellation. The existing literature treats cancel culture as a phenomenon that affects students uniformly, but both sociological theory and anecdotal evidence strongly suggest that the experience of being canceled, and its consequences, are significantly moderated by race, gender, socioeconomic status, sexual orientation, nationality, and institutional power (Alkhateeb & Bouherar, 2024). Alkhateeb and Bouherar (2024) gesture toward this gap in their examination of cancel culture in Arab academic contexts, acknowledging that political and social power structures shape who is vulnerable to cancellation and who retains the institutional protection to survive it, but systematic empirical data on differential vulnerability across demographic groups remains largely absent. Traversa et al. (2023) offer a partial corrective through their focus on marginalized group perspectives, finding that cancel culture is experienced as validating by those using it against powerful majorities but potentially weaponized against the same groups in different contexts, a paradox that urgently requires intersectional investigation.

A third underexplored dimension is the institutional and policy gap in higher education's response to cancel culture. Edling (2026) addresses the macro-level challenge of preserving academic freedom and moral citizenship education in higher education, but the existing literature provides limited guidance on what specific institutional interventions, counseling protocols, digital literacy curricula, restorative justice mechanisms, are effective in mitigating the psychological and academic consequences of cancel culture for students in different contexts. Van Grunsven and Marin (2024) propose digital skills and literacy courses as a mitigation strategy but acknowledge that no empirical evaluation of such interventions in cancel culture contexts has been published. Jochan and Banerjee (2021) note that survey respondents report altering their online behavior after experiencing digital shaming, but whether this behavioral change reflects adaptive coping or maladaptive self-censorship is not examined. Future research must therefore urgently address these gaps through longitudinal designs, intersectional sampling frameworks, and controlled evaluations of institutional responses, methodological investments that will enable the field to move from describing the problem to solving it.

Table 4: Summary of Identified Gaps in the Literature

Gap Category	Description	Why It Matters	Suggested Future Direction
Longitudinal Studies	No multi-year tracking of cancel culture's mental health and career outcomes	Cannot establish causality or recovery trajectories	Multi-wave cohort studies of university populations
Intersectional Analysis	Insufficient data on race, gender, sexuality as moderators of cancellation vulnerability	Hides differential burden and structural inequities	Demographically stratified sampling in future studies
Institutional Response Research	No empirical evaluation of digital literacy or restorative justice interventions	Cannot identify what works to protect students	Randomized controlled trials of campus interventions
Canceler Psychology	Minimal study of long-term mental health of cancel culture perpetrators	Incomplete picture of harm distribution	Longitudinal studies including all stakeholder roles
Non-Western & Global Contexts	Most research centers on Western, English-speaking universities	Limits generalizability and misses cultural variation	Cross-national comparative studies
Academic Performance Data	No systematic linkage between cancellation events and academic records	Cannot quantify the educational cost of cancel culture	Administrative data linkage with student welfare records

CONCLUSION

Cancel culture within digital student communities frequently operates less as a genuine tool for accountability and more as a cyberpsychological weapon that enforces rigid social conformity. Propelled by the online disinhibition effect, algorithmic echo chambers, and the developmental need for young adults to virtue-signal their belonging, digital ostracism inflicts severe psychological harm on its targets. By stripping context and demanding absolute moral purity, cancel culture weaponizes toxic shame. This leads to measurable mental health crises, including elevated rates of clinical depression, anxiety, and suicidal ideation, alongside profound academic sabotage. Because of the digital permanence of these platforms, targets face a lifelong reputational and psychological burden. Furthermore, the practice degrades the broader campus climate by inducing collective anxiety and self-censorship among bystanders, while subjecting active participants to moral grandstanding fatigue and hypervigilant paranoia.

Universities and educational institutions face a dual crisis driven by these cyberpsychological phenomena: a worsening student mental health emergency and a direct threat to the core epistemic function of academic inquiry. Traditional anti-bullying policies are insufficient because they often fail to capture the systemic, networked, and performative nature of digital ostracism. Institutions must recognize that cancel culture is not merely a series of isolated interpersonal conflicts, but a structural psychosocial hazard that actively suppresses academic freedom and intellectually diverse discourse. Failure to adapt institutional policies to

these realities abandons vulnerable students to digital vigilantism and replaces the university's mandate of open debate with dogmatic conformity and "safetyism."

To counteract the toxic echoes of cancel culture, educational institutions must proactively implement targeted digital empathy and digital citizenship curricula. These programs should directly address the socioethical challenges of online interaction and the psychological impact of public shaming. Furthermore, institutions should pivot away from purely punitive disciplinary models and establish restorative justice frameworks. These frameworks can mediate online conflicts by allowing for genuine accountability, contextualization, and community reintegration without permanent stigmatization. Finally, robust and specialized mental health support systems must be established to assist victims of online mobbing, equipping them with strategies to build shame resilience and navigate the trauma of context collapse and digital permanence.

To move beyond merely describing the phenomenon, future scholarship must address critical gaps in the literature. Longitudinal studies are urgently needed to track the long-term career, academic, and psychological impacts of collegiate cancellation. Additionally, intersectional research is required to understand how the vulnerabilities and consequences of cancel culture differ across varying demographic groups, including race, gender, and socioeconomic status. Lastly, empirical, controlled evaluations of institutional interventions, such as digital literacy initiatives and restorative justice mechanisms, must be conducted to identify evidence-based strategies capable of mitigating the harms of online mobbing and restoring psychological safety to digital student communities.

REFERENCES

- Alkhateeb, H., & Bouherar, S. (2024). Understanding cancel culture in higher education in the Arab world. *Higher Education Policy*, 37(2), 326–347. <https://doi.org/10.1057/s41307-023-00308-z>
- Chonge, H., Kiaritha, N., & Okapi, R. (2025). Cancel culture as a stressor: Implications for student's resilience and coping. *International Journal of Research and Innovation in Social Science (IJRISS)*, 9(11). <https://doi.org/10.47772/IJRISS.2025.9110006>
- Diether, E., Yi, S., Argyle, L. P., & Busby, E. C. (2024). The political psychology of cancel culture: Value framing or group identity? *Political Research Quarterly*, 77(4), 1130–1145. <https://doi.org/10.1177/10659129241261374>
- Edling, S. (2026). Challenges to moral citizenship education in Sweden higher education. *Moral Development and Citizenship Education*, 159. https://doi.org/10.1007/978-3-031-73961-3_9
- Jochan, G. M., & Banerjee, T. (2021). Shaming in the internet era: Evaluating the reintegrative function of shame in digital spaces. In *Shame 4.0: Investigating an emotion in digital worlds and the fourth industrial revolution* (pp. 431–454). Springer International Publishing. https://doi.org/10.1007/978-3-030-72884-0_21
- Julianti, S. (2023). Cancel culture: Cyberbullying on Twitter seen from the Space Transition Theory. *Jurnal Sosial Humaniora*, 14(2), 162–176. <https://doi.org/10.21111/jsh.v14i2.9727>
- Karim, F., Oyewande, A. A., Abdalla, L. F., Ehsanullah, R. C., & Khan, S. (2020). Social media use and its connection to mental health: A systematic review. *Cureus*, 12(6), e8627. <https://doi.org/10.7759/cureus.8627>
- Li, Y., Gao, X., & Li, H. (2022). The association between cyberbullying victimization and suicidal ideation among Chinese college students: The parallel mediating roles of core self-evaluation and depression. *Frontiers in Psychology*, 13, 892205. <https://doi.org/10.3389/fpsyg.2022.892205>

- Lokhande, G., & Natu, S. (2022). "You are cancelled": Emergence of cancel culture in the digital age. *IAHRW International Journal of Social Sciences Review*, 10(2), 215–220.
- Mueller, T. S. (2021). Blame, then shame? Psychological predictors in cancel culture behavior. *The Social Science Journal*, 1–14. <https://doi.org/10.1080/03623319.2021.1949617>
- Muir, S. R., Roberts, L. D., Sheridan, L., & Coleman, A. R. (2023). Examining the role of moral, emotional, behavioural, and personality factors in predicting online shaming. *PLoS One*, 18(3), e0279750. <https://doi.org/10.1371/journal.pone.0279750>
- Ng, E. (2022). *Cancel culture: A critical analysis*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-54436-3>
- Picarella, L. (2024). Intersections in the digital society: Cancel culture, fake news, and contemporary public discourse. *Frontiers in Sociology*, 9, 1376049. <https://doi.org/10.3389/fsoc.2024.1376049>
- Shannon, H., Bush, K., Villeneuve, P. J., Hellems, K. G., & Guimond, S. (2022). Problematic social media use in adolescents and young adults: Systematic review and meta-analysis. *JMIR Mental Health*, 9(4), e33450. <https://doi.org/10.2196/33450>
- Steinsbekk, S., Nesi, J., & Wichstrøm, L. (2023). Social media behaviors and symptoms of anxiety and depression: A four-wave cohort study from age 10–16 years. *Computers in Human Behavior*, 147, 107859. <https://doi.org/10.1016/j.chb.2023.107859>
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, 7(3), 321–326. <https://doi.org/10.1089/1094931041291295>
- Traversa, M., Tian, Y., & Wright, S. C. (2023). Cancel culture can be collectively validating for groups experiencing harm. *Frontiers in Psychology*, 14, 1181872. <https://doi.org/10.3389/fpsyg.2023.1181872>
- Ungerer, L. M. (2021). Cybershaming never rests: Suggestions for dealing with cybershaming in a digital culture. In *Shame 4.0: Investigating an emotion in digital worlds and the fourth industrial revolution* (pp. 555–573). Springer International Publishing. https://doi.org/10.1007/978-3-030-72884-0_27
- Van Grunsven, J., & Marin, L. (2024). Technosocial disruption, enactivism, & social media: On the overlooked risks of teenage cancel culture. *Technology in Society*, 78, 102602. <https://doi.org/10.1016/j.techsoc.2024.102602>
- Hossain, M. A., et al. (2023). Influence of cyber-victimization and other factors on depression and anxiety among university students in Bangladesh. *Frontiers in Psychiatry*, 14, 1272612. <https://doi.org/10.3389/fpsyg.2023.1272612>