

## Article

# Reshaping Digital Social Reality in the AI Era: A Data-Driven Analysis of University Students' Exposure to Digital Harassment in Emerging Countries

Mostafa Aboulmour Salem

Deanship of Development and Quality Assurance, King Faisal University, Al-Ahsa 31982, Saudi Arabia; masalem@kfu.edu.sa

## Abstract

Digital harassment is an increasing challenge in higher education, with implications for students' psychological well-being, perceived safety, and engagement in digital learning. As artificial intelligence (AI) increasingly mediates communication, visibility, and interaction across educational platforms, students' exposure to online harm is shaped not only by individual behaviour but also by algorithmically structured interaction environments. Understanding these conditions is essential for protecting student well-being and supporting sustainable participation in AI-enhanced learning. This study examines university students' exposure to digital harassment in AI-mediated learning environments using an expanded Unified Theory of Acceptance and Use of Technology (UTAUT) framework. Survey data were collected from 2185 students, including Saudi nationals and international students enrolled in Saudi Arabian universities, representing Saudi Arabia and 32 other developing and emerging countries (33 countries in total). The model analyses associations among technological literacy, cybersecurity awareness, social media engagement intensity, digital identity visibility, AI-mediated interactions, and cultural norms, while also accounting for disciplinary and cultural context differences. The results indicate that AI-mediated interactions are most strongly associated with exposure to digital harassment. Higher social media engagement, more restrictive cultural norms, and greater visibility of digital identity are associated with increased exposure, whereas technological literacy and cybersecurity awareness are associated with lower reported exposure. Furthermore, greater exposure to digital harassment is linked to poorer mental health outcomes and reduced continuity in e-learning participation. Overall, the findings suggest that digital harassment in AI-driven educational settings is a structural sociotechnical issue associated with greater embeddedness in algorithmically mediated learning environments, rather than an isolated behavioural issue. The study highlights the need for responsible AI governance, enhanced digital literacy education, and culturally responsive institutional policies to support inclusive and sustainable higher education.

Academic Editors: Felice Addeo and Gabriella Punziano

Received: 7 December 2025

Revised: 5 February 2026

Accepted: 12 February 2026

Published: 21 February 2026

**Copyright:** © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

**Keywords:** artificial intelligence (AI); algorithmic governance; digital harassment; digital identity visibility; technological literacy; cybersecurity awareness; social media engagement; AI-mediated interactions; digital inequalities; mental health; E-learning continuity; data-driven social research; sustainable development goals (SDGs)

## 1. Introduction

Digital harassment has become a growing and deeply consequential challenge within contemporary higher education. As universities increasingly rely on digital and AI-enhanced platforms for teaching, communication, and assessment, students' academic and social lives are now embedded within data-driven environments that shape visibility, interaction, and social participation [1,2].

Beyond being a peripheral online is associated with higher levels of cyberbullying, and related forms of digital harassment are widespread and severe: global evidence indicates that approximately 14–57% of students have experienced cyberbullying, depending on context and measurement approaches [3]. Significantly, cyberbullying is strongly associated with severe mental health outcomes, including anxiety, depression, emotional distress, and social withdrawal [4–7].

Large-scale epidemiological research further shows that students exposed to cyberbullying are more than four times as likely to report suicidal thoughts and suicide attempts compared with non-victims, even after controlling for social and family factors [3]. In parallel, digital harassment has been shown to undermine educational participation, contributing to reduced concentration, avoidance of online platforms, absenteeism, and disengagement from learning activities [8,9].

These findings demonstrate that digital harassment within AI-mediated educational environments is not merely a behavioural inconvenience but a serious threat to student safety, mental health, and educational continuity, raising urgent concerns about equity, well-being, and institutional responsibility in data-driven higher education [3].

AI-driven systems—such as algorithmic recommendations, automated content ranking, behavioural profiling, and AI-supported communication tools—play an increasingly active role in structuring online interaction [1]. Rather than serving as neutral infrastructures, these systems influence what students see, whom they interact with, and how long content or interactions remain visible [2].

Consequently, digital environments do not merely host social behaviour; they actively organise it. Within higher education, this shift has profound implications for students' exposure to online harm and their capacity to engage safely and sustainably in learning [3,4]. Although digital technologies have expanded access to education and collaboration, they have also intensified students' exposure to digital harassment [4–6].

Cyberbullying, cyber hate, sexual extortion, impersonation, and digital stalking are no longer marginal phenomena but persistent issues associated with higher levels that undermine students' psychological well-being, sense of safety, academic engagement, and willingness to remain active in online learning environments [7–10]. Empirical research consistently links exposure to such harms with anxiety, emotional exhaustion, reduced concentration, lower academic performance, and, in severe cases, withdrawal from digital learning communities and diminished educational motivation [7,11].

Crucially, digital harassment in contemporary learning environments cannot be understood solely as the product of individual interpersonal behaviour. In AI-mediated contexts, algorithmic systems may unintentionally reshape exposure conditions by amplifying visibility, facilitating repeated or unsolicited contact, and extending the persistence and circulation of harmful content [12,13]. This sociotechnical dynamic aligns with scholarship in digital sociology and algorithmic governance, which emphasises that technological infrastructures actively configure social relations rather than merely reflecting them [3,14].

Accordingly, understanding digital harassment requires moving beyond individual-level explanations to examine how algorithmic mediation restructures the conditions under which harm becomes more likely, sustained, or difficult to escape [15]. These challenges are particularly acute in developing and emerging countries, where disparities in

technological literacy, uneven access to digital resources, and culturally embedded norms influence both vulnerability and reporting behaviour [16,17].

In such contexts, digital inequalities translate into stratified patterns of exposure, reinforcing existing social divides and producing uneven access to safe and inclusive online learning spaces [11]. Universities operating within these environments therefore face heightened ethical and institutional responsibilities, especially under international commitments such as the Sustainable Development Goals—most notably SDG 4 (Quality Education) and SDG 5 (Gender Equality)—which call for safe, inclusive, and non-discriminatory educational settings [18,19].

Despite a growing body of research on cyberbullying and online harassment (e.g., [20–22]), significant gaps remain. Much of the existing literature focuses on general social media contexts or individual behavioural factors associated with higher levels, often overlooking how technological literacy, social media engagement intensity, digital identity visibility, and cultural norms interact within AI-driven educational environments. Moreover, sociotechnical dimensions—such as algorithmic curation, platform governance, automated moderation, and data-driven interaction patterns—remain underexamined, particularly in higher education settings where digital participation is frequently compulsory rather than optional.

In addition, relatively few studies explicitly distinguish between established behavioural factors associated with higher levels of factors and the amplification associations introduced by AI-mediated interactions. As AI becomes increasingly embedded in learning management systems, communication platforms, and academic support tools, understanding how algorithmic mediation reshapes exposure to digital harassment is essential for developing associationive, context-sensitive prevention and intervention strategies.

To address these gaps, the present study examines university students' exposure to digital harassment within AI-enhanced learning environments using an expanded Unified Theory of Acceptance and Use of Technology (UTAUT) framework. The study investigates five key associated factors—technological literacy and cybersecurity awareness, social media engagement intensity, digital identity visibility, AI-mediated interactions, and cultural norms and social expectations—and analyses how these factors jointly shape vulnerability to online harassment. It further examines the consequences of exposure for students' mental health and continuity in e-learning, while exploring the moderating roles of academic specialisation and cultural context to capture disciplinary and socio-cultural variation associated with higher levels of patterns.

Accordingly, the inferential population of this study comprises both Saudi nationals and international students enrolled in Saudi Arabian universities, allowing examination of AI-mediated experiences within a shared institutional and technological environment. In this study, developing and emerging countries are defined as low-, lower-middle-, or upper-middle-income economies according to the World Bank's income classification. Country-of-origin classifications are used solely for descriptive and analytical purposes. References to the Human Development Index (HDI) and the IMF World Economic Outlook (WEO) are included only for contextual background and do not affect sample inclusion, which is governed exclusively by World Bank income categories.

Saudi Arabia provides a shared host-country educational environment with relatively uniform digital infrastructure and institutional governance. At the same time, students' countries of origin reflect diverse sociocultural and technological conditions relevant to the dynamics of digital harassment. By focusing on AI-mediated educational ecosystems, this study advances existing research beyond general online interaction contexts. It conceptualises digital harassment as a structural sociotechnical phenomenon associated with higher levels of vulnerability, inequality, and student well-being within contemporary higher education, shaped by behavioural practices, cultural constraints, and

algorithmic mediation, offering a more comprehensive understanding of how digital societies are reshaping vulnerability, inequality, and student well-being within contemporary higher education.

## 2. Literature Review and Research Hypotheses

### 2.1. Associated Factors of Exposure to Digital Harassment in AI-Mediated Environments

The expansion of AI-enhanced platforms has reshaped social interaction in higher education by embedding student behaviour within algorithmically governed environments that structure visibility, identity performance, and interpersonal exposure [23,24]. As a result, digital harassment emerges not only from interpersonal conflict but also from sociotechnical conditions created by data-driven mechanisms that influence whom students encounter, how they are seen, and how interactions persist online [25–27].

For conceptual clarity and to minimise construct overlaps, the core associated factors examined in this study are defined as follows. Technological Literacy and Cybersecurity Awareness (TL&CA) refers to the extent to which students possess practical digital safety knowledge and skills, including managing privacy settings, protecting accounts and devices, recognising common cyber threats (e.g., phishing), and applying secure behaviours when using academic and social platforms [27,28].

Social Media Engagement Intensity (SMEI) captures the frequency and habitual intensity of students' social media use—such as checking, browsing, posting, and interacting during daily routines—which is associated with greater opportunities for exposure in online spaces [24,29]. Digital Identity Visibility (DIV) denotes the degree to which a student's identity is publicly discoverable and informationally exposed through profile openness, self-disclosure, identifiable images, and public-facing activity, factors associated with greater audience reach and potential targeting [30,31].

AI-Mediated Interactions (AIMI) represent the extent to which students' learning and social interactions are shaped by AI-enabled features—such as recommendation and ranking systems, automated prompts, suggested connections, and chatbot-assisted communication—that structure what content is encountered and with whom interactions occur [30,32]. This construct captures algorithmic mediation of interaction contexts rather than direct AI-enabled harassment modalities (e.g., deepfakes or impersonation), unless explicitly measured by the items. Cultural Norms and Social Expectations (CNSE) reflect perceived socio-cultural expectations that shape how students interpret, tolerate, report, or respond to digital harassment, including stigma, honour-related concerns, and the normalisation of online aggression, all of which influence vulnerability and help-seeking behaviour [27,28].

Exposure to Digital Harassment (EDH) refers to the frequency of experiencing harmful, threatening, humiliating, or invasive online behaviours—including general harassment, hate, sexualised harassment or sextortion, and digital stalking or monitoring—across digital platforms [4,15,33]. Mental Health Impact (MHI) captures perceived psychological strain associated with harassment exposure, such as distress, anxiety, fear, reduced well-being, and impaired concentration, consistent with evidence linking cybervictimisation to adverse mental health outcomes [8,34]. E-Learning Continuity (ELC) reflects sustained participation and the intention to continue engaging in online learning activities, including attendance, interaction, and persistence, in line with continuance-intention and e-learning research [35,36].

Additionally, recent empirical studies show that higher levels of digital literacy and cybersecurity awareness typically reduce vulnerability to digital harassment by enabling students to actively manage privacy settings, recognise suspicious activity, and adopt safer online practices [28,37,38]. In contrast, low digital literacy is consistently associated with greater online exposure, which in turn is associated with higher levels of harmful

interactions [3,4,13]. Furthermore, intensive social media engagement and discernible digital identities have been consistently linked to increased harassment, as frequent posting, commenting, and public profiles attract more attention, including from potential aggressors [14,22].

Likewise, AI-mediated interactions further alter this landscape: algorithmically curated feeds, automated messaging, and recommendation systems may unintentionally expose students to harmful content or malicious users, particularly when safety mechanisms and platform governance are weak [32,39]. Cultural norms and social expectations—such as stigma around reporting [40], honour-based concerns [41], honour-based concerns or the normalisation of online aggression [42]—may further intensify the issue, associated with higher levels by discouraging disclosure and enabling harmful behaviours to persist. Based on this literature and the expanded Unified Theory of Acceptance and Use of Technology (UTAUT)-based framework, the following hypotheses are proposed:

**H1:** *TL&CA is negatively associated with exposure to EDH.*

**H2:** *SMEI is positively associated with exposure to EDH.*

**H3:** *DIV is positively associated with exposure to EDH.*

**H4:** *AIMI is positively associated with exposure to EDH.*

**H5:** *CNSE is positively associated with exposure to EDH.*

This study adopts an expanded UTAUT as a behavioural and contextual framework to examine patterns of exposure to digital harassment, rather than as a statistical theory of victimisation. Although UTAUT was initially developed to explain technology acceptance and use, its core constructs—facilitating conditions, effort expectancy, and social influence—provide a valuable lens for understanding how students navigate digital platforms, become visible online, and encounter issues associated with higher levels of AI-mediated environments.

From this perspective, exposure to digital harassment is shaped not only by harmful behaviour from other users but also by how platforms are used, how visible students are within them, and how algorithmic systems organise content and interaction opportunities. Accordingly, EDH is conceptualised as an experiential and situational outcome arising from the interaction of behavioural, technological, and socio-cultural conditions, rather than as the direct result of any single statistical factor.

## 2.2. Exposure, Mental Health, E-Learning Continuity, and Moderating Contexts

Prior research consistently shows that exposure to digital harassment, including cyberbullying [4], cyber hate [7], sexualised harassment [43], and digital stalking has been linked to heightened anxiety [3], emotional exhaustion [14], depressive symptoms [33], and a diminished sense of safety [13], all of which contribute to profound psychological and academic consequences for university students, particularly in AI-mediated learning environments where digital participation is essential for course engagement [32].

Furthermore, psychological burdens reduce students' motivation, impair concentration, and weaken their willingness to engage in online learning [5]. Sustained exposure can also disrupt e-learning continuity, leading students to withdraw from digital platforms, avoid synchronous sessions, or limit participation in virtual classrooms [1,44]. In line with this evidence, the study proposes that exposure to digital harassment will negatively influence both mental health and e-learning continuity:

**H6:** EDH is negatively associated with MHI.

**H7:** EDH is negatively associated with ELC.

However, vulnerability to digital harassment is unlikely to be uniform across all students. As well, the academic specialisation may influence how behaviourally associated factors such as technological literacy, social media engagement, and digital identity visibility translate into exposure, given disciplinary differences in digital workload, communication norms, and platform usage [6,16]. Likewise, cultural context is expected to shape how AI-mediated interactions and cultural norms affect exposure, reflecting cross-cultural variation in attitudes toward AI, privacy, reporting behaviour, and online aggression [45,46]. Therefore, the model anticipates that academic specialisation and cultural context factors will moderate several association relationships:

**H8:** AS moderates the association between TL&CA and EDH.

**H9:** moderates the association between SMEI and EDH.

**H10:** AS moderates the association between DIV and EDH.

Similarly, cultural context is expected to moderate relationships involving AI-driven interactions and cultural norms:

**H11:** CC moderates the association between AIMI and EDH.

**H12:** CC moderates the association between CNSE and EDH.

Together, these moderating pathways reveal a nuanced pattern in which digital is associated with higher levels shaped by behavioural factors, academic contexts, and socio-cultural conditions within AI-driven higher education systems. In this study, AI-mediated interactions are conceptualised not as a distinct form of harassment, but as algorithmic affordances that reshape exposure conditions—such as visibility, reach, and persistence—within digital learning environments.

### 2.3. Research Gaps

The reviewed literature consistently demonstrates that digital harassment negatively affects students' well-being and participation in online environments. However, several important gaps remain. First, most studies conceptualise harassment primarily as an interpersonal or behavioural issue, with limited attention to AI-mediated interaction mechanisms—such as algorithmic recommendations, visibility amplification, and automated communication—that actively shape exposure conditions in contemporary learning platforms. Second, existing research rarely integrates behavioural, technological, and cultural associated factors within a single explanatory framework to model exposure in higher education contexts where digital participation is often compulsory. Third, downstream academic consequences, particularly e-learning continuity, remain underexplored despite the central role of AI-enhanced platforms in assessment, communication, and learning persistence. These limitations are especially evident in developing and emerging countries, where rapid digital adoption often outpaces institutional, regulatory, and pedagogical safeguards. (See Table 1).

**Table 1.** Summary of Key Literature on Digital Harassment in AI-Mediated Higher Education.

| Study                                     | Context/Sample               | Focus                                    | Key Findings                                                         | Identified Gap                                                                   |
|-------------------------------------------|------------------------------|------------------------------------------|----------------------------------------------------------------------|----------------------------------------------------------------------------------|
| Myers and Cowie (2019) [11].              | Education (school–HE)        | Cyberbullying across educational stages  | Cyberbullying undermines well-being and participation                | Does not address AI-mediated interaction structures or algorithmic exposure      |
| Oksanen et al. (2022) [13].               | Academic online environments | Online hate and harassment               | Digital platforms intensify harassment associated with higher levels | Lacks student-level modelling and AI-enabled learning contexts                   |
| Wang et al. (2022) [39].                  | Online learning systems      | AI-mediated interaction design           | AI reshapes interaction and governance                               | Does not model harassment exposure or psychosocial outcomes                      |
| Elrayah and Jamil (2023) [37].            | Cybersecurity behaviour      | Digital literacy and awareness           | Higher literacy improves online safety                               | Limited linkage to harassment exposure and educational continuity                |
| Johansen and Tjørnhøj-Thomsen (2024) [43] | Youth experiences            | Digital sexual harassment                | Strong psychosocial consequences                                     | Limited examination of algorithmic amplification mechanisms                      |
| Swain (2024) [10].                        | Higher education             | Online harms and institutional responses | Harassment requires systemic HE policies                             | Lacks an integrated framework linking associated factors, exposure, and outcomes |
| Eslen-Ziya et al. (2024) [22]             | Online visibility            | Identity exposure and harassment         | Visibility is associated with higher levels of vulnerability         | Focuses on academics; limited relevance to student AI-learning contexts          |
| Shaikh et al. (2025) [20]                 | Quantitative SEM             | Correlates of cyber-harassment           | Multi-factor models explain the association with higher levels       | AI treated implicitly; limited focus on AI-mediated interactions                 |

Building on these gaps, the present study advances literature by conceptualising digital harassment in AI-enhanced higher education as a structural sociotechnical phenomenon associated with higher levels of, rather than solely an outcome of, individual misconduct. Drawing on an expanded Unified Theory of Acceptance and Use of Technology (UTAUT), the study integrates behavioural factors (social media engagement and identity visibility), technological conditions (AI-mediated interactions and cybersecurity awareness), and sociocultural influences (cultural norms and social expectations) to explain students' exposure to digital harassment. It further examines the consequences of exposure for mental health and e-learning continuity, while testing the moderating associations of academic specialisation and cultural context. Based on this framework, the following hypotheses are proposed.

### 3. The Research Methods

#### 3.1. Research Population and Sample

The World Economic Outlook (WEO) classifies countries into two broad categories: advanced economies (e.g., the United States, the United Kingdom, Italy, and Japan) and emerging and developing economies (e.g., Saudi Arabia, Egypt, Qatar, and Nigeria) [47]. In the present study, WEO terminology is used solely for descriptive contextualisation of global economic groupings. Country inclusion and analytical classification are governed exclusively by World Bank income categories, while Human Development Index (HDI) levels and IMF classifications are referenced only for contextual interpretation.

The Kingdom of Saudi Arabia serves as a regional hub for multicultural higher education and digitally mediated learning environments. During the 2024–2025 academic

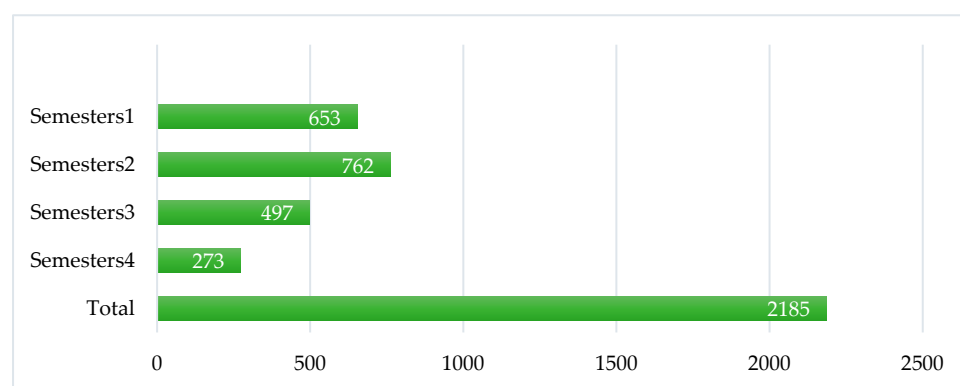
year, Saudi universities enrolled approximately 98,453 international students, in addition to Saudi nationals, representing a wide range of national and cultural backgrounds [48].

Accordingly, this study focused on Saudi and international students enrolled in Saudi Arabian universities. Saudi Arabia was deliberately selected as a shared host country to ensure consistency in key institutional and technological conditions, including digital infrastructure, learning management systems (LMS), platform governance, and institutional policies regulating online learning and digital interaction. This design enables examination of exposure to digital harassment within a relatively uniform regulatory and technological environment, while preserving meaningful cross-national variation associated with students' countries of origin.

Universities were included in the sample if they met the following criteria: (i) provision of AI-enhanced or digitally mediated learning environments, such as institutionally managed LMS platforms, online assessments, or AI-supported educational tools; (ii) enrolment of a substantial population of Saudi nationals and international students studying within the same Saudi Arabian higher-education context; and (iii) reliance on formally governed, institutionally managed digital systems rather than informal or ad hoc platforms for teaching, learning, and academic communication. These criteria ensured that all participating students were embedded in comparable AI-mediated educational ecosystems aligned with the study's objectives.

The final analytical sample comprised 2185 students (991 males and 1194 females) originating from Saudi Arabia and 32 other developing and emerging countries (33 countries in total; see Appendix A). This sampling context provides a unique opportunity to examine digital interaction patterns within a unified host-country environment while simultaneously capturing sociocultural heterogeneity linked to students' national backgrounds. Such variation is particularly relevant because exposure to digital harassment is shaped by sociotechnical systems, culturally embedded norms, and differing expectations regarding online behaviour and reporting. Consequently, the use of a culturally heterogeneous sample enhances the study's capacity to identify differential patterns of vulnerability across social and cultural groups.

Data were collected over an extended period, from January 2024 to November 2025, to facilitate recruitment of a large, culturally diverse sample across multiple institutions. Although data collection spanned several academic semesters, the study employed a cross-sectional design, as each participant completed the survey only once and no repeated measurements were obtained (see Figure 1).



**Figure 1.** Data collection semesters.

Consequently, the dataset captures a single observational snapshot of students' experiences and perceptions rather than longitudinal change. (The dataset is available at the following link: <https://2u.pw/mRgWI> (accessed on 18 January 2026))

To assess potential temporal heterogeneity arising from the extended data collection window, additional robustness checks were conducted by comparing key latent constructs across different collection phases. No statistically significant differences were observed across periods (all  $p > 0.05$ ), indicating that the estimated relationships are stable and not driven by time-varying contextual conditions.

To ensure sufficient statistical power for the complex structural model, which includes moderation associations, power analysis was conducted using inverse-square-root and gamma-exponential methods, indicating a minimum sample size of approximately 150 for medium-sized associations, with  $\alpha = 0.05$  and power = 0.80 [47]. The sample size achieved of 2185 participants exceeds the required number. Post hoc power analysis, using the smallest observed path coefficient ( $\beta = 0.083$ ,  $p = 0.219$ ), confirmed a power greater than 0.95, verifying the sample's adequacy for detecting minor to moderate associations, including moderator interactions [48].

### 3.2. Moderating Constructs

In this study, Cultural Context (CC) is operationalised as a macro-cultural grouping variable based on respondents' country of origin (demographic item). It is used exclusively for moderation and heterogeneity assessment [49]. CC does not represent an individual-level scale score and is conceptually distinct from Cultural Norms and Social Expectations (CNSE), which captures respondents' personal perceptions of reporting stigma, honour-related concerns, and social expectations surrounding digital harassment [50]. CC was coded into three regional groups: Middle East, Africa, and Eastern Asia, with percentages calculated relative to the total sample ( $N = 2185$ ) (see Table A1).

Moderation by CC was examined using interaction terms in the primary PLS-SEM specification. To enhance the interpretability and robustness of cross-cultural comparisons, complementary group-based analyses were conducted using PLS multigroup analysis (MGA) after establishing measurement invariance [48]. Measurement invariance was assessed using the MICOM procedure, including configural, compositional, and composite-mean and variance equality. Cross-group differences in structural path estimates were interpreted only after MICOM requirements were satisfied, using permutation- and bootstrap-based MGA procedures implemented in SmartPLS [51].

Therefore, because respondents may be nested within universities and national contexts, ignoring within-cluster correlation could inflate Type I error rates. Although full multilevel SEM is not always feasible in variance-based SEM, this potential bias was explicitly considered. Accordingly, robust analytical strategies commonly recommended for clustered data were adopted to ensure that standard errors and statistical inferences were not unduly affected by the sample's nested structure.

Furthermore, Academic Specialisation (AS) was operationalised as a categorical moderator based on students' self-reported field of study. To ensure theoretical coherence and adequate group sizes for moderation analysis, individual majors were aggregated into three analytically meaningful clusters commonly used in higher-education research: (i) STEM (Sciences, Engineering, Computer Sciences, and Health-Related Disciplines), (ii) Social Sciences and Humanities (Education, Law, Arts, and Related fields), and (iii) Business and Applied Studies (Business Administration, and Related fields).

Moreover, these categories reflect systematic differences in digital workload, online visibility, and platform-mediated interaction intensity reported in prior digital-behaviour research. For PLS-SEM moderation, AS was dummy-coded using a binary contrast, with Social Sciences and Humanities as the reference group. This coding strategy enables direct interpretation of interaction associations as differences in structural path strengths relative to the reference category.

In addition, moderation hypotheses (H8–H12) were tested using the product-indicator interaction approach implemented in SmartPLS. All correlate constructs in interaction terms were mean-centred prior to construction to reduce multicollinearity and facilitate interpretation of conditional associations. Interaction terms were entered into the structural model only after the significance and stability of the corresponding main associations had been established. Likewise, the statistical significance of interaction paths was assessed using bootstrapping with 5000 resamples. In addition to path coefficients and  $p$ -values, interaction association sizes ( $f^2$ ) were computed to evaluate the substantive relevance of the moderation associations, following established PLS-SEM reporting guidelines [52].

For all statistically significant interaction associations, simple slope analyses were conducted to visualise and interpret how the relationship between each correlation and Exposure to Digital Harassment (EDH) varies across levels of the moderator. Continuous correlates were plotted at  $\pm 1$  standard deviation from the mean, while categorical moderators (academic specialisation and cultural context) were visualised using group-specific conditional associations derived from the interaction model. Hence, simple-slope plots were generated directly in SmartPLS using the interaction-plot function, and the latent-variable scores were cross-validated. (See Figures 2 and 3)

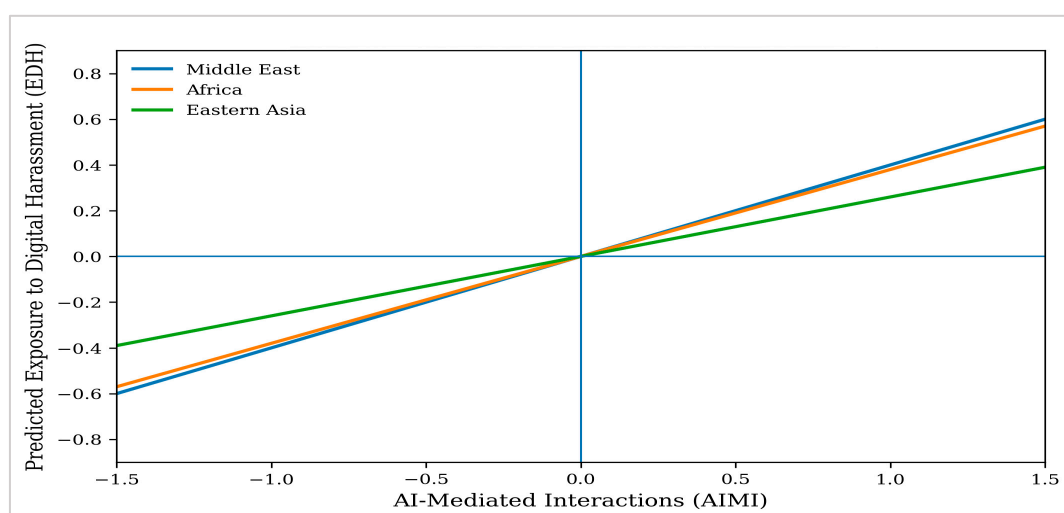


Figure 2. Cultural Context.

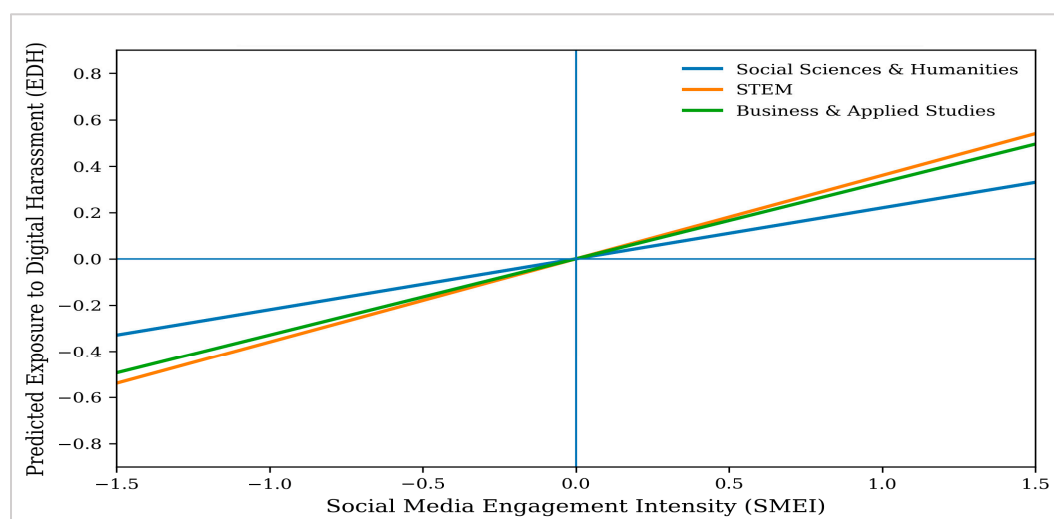


Figure 3. Academic Specialisation.

### 3.3. Data Collection and Instrument Design

Data were collected using an anonymous online questionnaire designed to minimise social desirability bias and address the sensitive nature of reporting digital harassment. Participation was entirely voluntary, and respondents were assured that all responses would remain confidential. The survey link was distributed through official institutional email channels, with an invitation emphasising the study's relevance to improving digital safety and informing the future of AI-enhanced higher education.

The questionnaire comprised three sections; the first collected demographic information, including gender, academic specialisation, and cultural context. The second section presented an informed consent statement outlining the study's purpose, ethical safeguards, and participants' rights. The third section included Likert-scale items measuring the constructs of the proposed theoretical model.

All measurement items were adapted from established and validated instruments to ensure conceptual accuracy and reliability. Technological Literacy and Cybersecurity Awareness (TL&CA) was measured using five items (TL&CA1–TL&CA5) adapted from [27,28,37]. Social Media Engagement Intensity (SMEI) was measured using five items (SMEI1–SMEI5) adapted from [29,39]. Digital Identity Visibility (DIV) was evaluated using three items (DIV1–DIV3) as described in [22,29,30,53]. AI-Mediated Interactions (AIMI) was measured using four items (AIMI1–AIMI4) derived from [4,30,32,54]. Cultural Norms and Social Expectations (CNSE) was assessed using five items (CNSE1–CNSE5) adapted from [1,9,45].

Exposure to Digital Harassment (EDH) and Mental Health Impact (MHI) were measured using items informed by [3–6], comprising five EDH items (EDH1–EDH5) and four MHI items (MHI1–MHI4) [15,33,45]. Mental Health Impact (MHI) was measured using four items (MHI1–MHI4) assessing psychological distress, anxiety, reduced sense of safety, and impaired concentration resulting from harassment exposure. MHI was operationalised as a negatively valenced construct, with higher scores indicating poorer mental health outcomes. E-Learning Continuity (ELC) was assessed using four items (ELC1–ELC4) adapted from [16,20,55], capturing sustained participation and persistence in online learning activities.

To ensure cultural and linguistic appropriateness for the Saudi Arabian higher education context, the instrument underwent a rigorous localisation and validation process. Thirteen bilingual experts in higher education and information technology evaluated all items for content validity, cultural relevance, and conceptual clarity. In addition, ten university students reviewed the Arabic and English versions to assess readability and contextual suitability. All items were subjected to a forward–backward translation procedure verified by a bilingual committee to ensure linguistic and conceptual equivalence. Minor wording refinements were made to enhance clarity and contextual precision. A pilot study with 23 participants confirmed satisfactory reliability across all constructs [48].

All constructs were modelled as reflective latent variables and analysed using Partial Least Squares Structural Equation Modelling (PLS-SEM). Exposure to Digital Harassment (EDH) was specified as a reflective construct, consistent with measurement theory, when indicators represent correlated manifestations of an underlying experiential condition [56–58]. In this study, EDH is conceptualised as a frequency-based experiential outcome reflecting the intensity of exposure to harmful online behaviours across AI-mediated educational and social platforms. Accordingly, the EDH indicators are treated as parallel manifestations of a common latent exposure dimension, rather than as independent or formative components.

Furthermore, reflective modelling is appropriate when indicators are expected to be correlated, conceptually interchangeable manifestations of the same phenomenon, and are caused by the underlying latent construct rather than forming it. In contrast, formative

modelling is suitable when indicators represent non-substitutable facets whose omission would alter the construct's meaning [56,59,60].

Given the theoretical framing of EDH as an underlying exposure intensity in AI-mediated digital environments, the reflective specification aligns with measurement theory. It supports the interpretation of EDH as a unified experiential outcome.

Empirically, the reflective EDH model demonstrated strong psychometric properties consistent with expectations for reflective measurement, including high and relatively homogeneous indicator loadings and strong internal consistency. These results support the interpretation that the EDH items capture a standard latent exposure dimension. While EDH encompasses multiple forms of harassment, the intent is not to treat these as interchangeable social phenomena, but as co-occurring manifestations of a shared exposure condition that is appropriately modelled as a single latent construct.

### 3.4. Common Method Variance (CMV) Concerns

Common Method Variance (CMV) is a frequent concern in social science research, particularly when independent and dependent variables are self-reported by the same participants [61]. Such bias may inflate relationships among variables and affect the explanatory power and validity of the structural model [61].

To mitigate this, the study followed the recommendations of Podsakoff et al. [51] and implemented both procedural and statistical remedies. Procedurally, the questionnaire was carefully designed to minimise response bias by balancing item ordering and avoiding leading or patterned sequences. Dependent variables were deliberately positioned at later stages of the survey to reduce priming associations and mitigate respondents' tendency to infer relationships across items [62]. The questionnaire's overall length was also monitored to reduce fatigue-related bias.

Statistically, CMV was first assessed using Harman's single-factor test. The results showed that the first unrotated factor explained 41% of the total variance—well below the 50% threshold—indicating that CMV was unlikely to pose a significant threat to the study's findings. To further reinforce this conclusion, additional, more robust statistical techniques were employed. A marker-variable technique using an unrelated construct (attitude toward SDGs) revealed that the marker accounted for less than 2.5% of shared variance with the primary constructs, signalling no inflation of correlations. A latent method factor model was also evaluated using confirmatory factor analysis (CFA), and the model fit did not improve relative to the baseline model ( $\Delta\text{CFI} = 0.004$ ;  $\Delta\text{RMSEA} = 0.003$ ), confirming the absence of substantial method bias.

Data integrity checks were implemented throughout the collection period. Automated validation screened for straight-lining, extreme outliers, and duplicate IP submissions. Responses with completion times below 40% of the median duration were excluded ( $n = 11$ ), and no duplicate entries were retained. Enumerator-assisted completions were monitored through independent logs to ensure adherence to standardised administration procedures. Missing data were minimal (below 2% across all items), and Little's MCAR test was not statistically significant ( $\chi^2 = 14.82$ ,  $p = 0.26$ ), supporting the assumption that missingness was random. Because of the low proportion, missing values were imputed using expectation–maximisation (EM) prior to PLS-SEM analysis.

To avoid order associations, questionnaire sections were counterbalanced during pilot testing, and no significant mean differences were observed between split forms ( $p > 0.05$ ). The final version adopted the empirically validated sequence, beginning with demographic items and concluding with construct measures, ensuring a logical flow and sustained respondent engagement. Collectively, these procedural and statistical safeguards confirm that CMV, response bias, and order associations did not pose significant threats to the validity of the study's data or its structural model estimates.

### 3.5. Hypothesis Examining Approach

A structured, sequential analytical strategy was adopted to examine the hypothesised relationships within the expanded UTAUT-based framework. The analysis proceeded in two phases. First, the direct associations of the five associated factors—Technological Literacy and Cybersecurity Awareness (TL&CA), Social Media Engagement Intensity (SMEI), Digital Identity Visibility (DIV), AI-Mediated Interactions (AIMI), and Cultural Norms and Social Expectations (CNSE)—on Exposure to Digital Harassment (EDH) were evaluated (H1–H5). This step isolated the primary behavioural, technological, and cultural correlations of exposure, controlling for moderating variables. Consistent with methodological guidelines, moderation terms were excluded during the examination of main associations to avoid confounding and multicollinearity between direct and interaction terms.

In the second phase, the outcome pathways (H6–H7) and moderator hypotheses (H8–H12) were analysed. Exposure to Digital Harassment (EDH) was tested for its correlation with Mental Health Impact (MHI) and E-Learning Continuity (ELC), with psychological and academic consequences assessed. Academic Specialisation (AS) was modelled as a moderator of the relationships between TL&CA, SMEI, and DIV with EDH (H8–H10), while Cultural Context (CC) was tested as a moderator of the relationships between AIMI, CNSE, and EDH (H11–H12). Interaction terms were mean-centred and entered the model in accordance with recommended PLS-SEM moderation procedures.

Furthermore, the measurement model was first evaluated to confirm psychometric adequacy. Indicator reliability was assessed using outer loadings, and internal consistency was assessed using Cronbach's alpha and Composite Reliability. Convergent validity was examined via the Average Variance Extracted (AVE) [52]. Discriminant validity was assessed using both the Fornell–Larcker criterion and HTMT ratios, and all constructs met established thresholds [47]. Additionally, moderation hypotheses were tested by examining the significance of interaction terms, calculating moderation association sizes ( $f^2$ ), and interpreting simple slope visualisations for significant interactions. All structural analyses were conducted in SmartPLS 4, chosen for its robustness in modelling complex frameworks with multiple moderators, non-normal data, and reflective constructs [63].

Throughout the hypothesis testing and reporting of structural relationships, the term “association” is used solely as a statistical descriptor of estimated associations. Owing to the study's cross-sectional, observational design, statistical inference is neither assumed nor implied.

Likewise, to minimise common method bias, both procedural and statistical remedies were implemented. Anonymity, randomisation of item sequences, and separation of correlated and outcome items were applied during instrument design. Statistically, Harman's single-factor test indicated no general factor, full collinearity VIFs were below the recommended 3.3 threshold, and a latent method factor test revealed minimal loading shifts ( $<0.20$ ). Together, these results suggest that standard method variance was not a substantive threat to the validity of the findings.

### 3.6. Ethical Approvals

Before initiating data collection, formal institutional approval was obtained. Furthermore, this procedure certified that our employed methods were consistent with the institutional criteria and the ethical concerns outlined in the Declaration of Helsinki [64]. Additionally, several measures were implemented to protect participants' rights. Participation was entirely voluntary and free from coercion, with written informed consent obtained from each respondent. Participants were assured of their right to withdraw at any time without providing a reason. All data were anonymised to safeguard confidentiality. Respondents were informed that their answers would remain anonymous, securely stored

on encrypted institutional servers, and used solely for academic research purposes. No identifiable information was collected.

### 3.7. Software, Versions, and Reproducibility Details

All analyses were conducted using the following software and settings to support reproducibility. SmartPLS 4.0 (SmartPLS GmbH, Bönningstedt, Germany) was used for the primary PLS-SEM analysis, moderation modelling, bootstrapping with 5000 subsamples, MICOM measurement invariance testing, and simple-slope visualisation. IBM SPSS Statistics 29 was used for data screening, descriptive statistics, missing-data diagnostics, and preliminary robustness checks.

AMOS 29 was used to run a confirmatory factor analysis (CFA) with maximum likelihood estimation to assess standard method variance. Key analytical settings included the path weighting scheme, mean-centring of indicators for interaction terms, bias-corrected bootstrap confidence intervals, and two-tailed significance testing at  $\alpha = 0.05$ . These details are reported to enable independent replication of the whole analytical workflow.

## 4. Data Analysis Methods and Results

To evaluate the hypothesised model, Partial Least Squares Structural Equation Modelling (PLS-SEM) was employed as the primary analytical technique. This choice reflects the study's conceptualisation of digital harassment as a multidimensional sociotechnical phenomenon, shaped by behavioural, technological, and cultural associated factors operating within algorithmically mediated environments [62].

Contrasting the covariance-based SEM (CB-SEM), which typically requires large sample sizes and strict adherence to multivariate normality, PLS-SEM can accommodate smaller samples and does not rely on the normality assumption [61]. The data analysis was conducted employing SmartPLS (version 4) [51]. To guarantee robustness of the findings, a bootstrapping process with 5000 subsamples was performed with a reflective measurement procedure [65]. Following the recommendation from Hair et al. (2019) [52], the model assessment process was performed in two successive stages.

Table 2 presents the assessment of the measurement model, which demonstrates strong reliability and convergent validity across all constructs. All indicator loadings exceeded the recommended threshold of 0.70, confirming that each item reliably represents its corresponding latent variable. For instance, loadings for Technological Literacy and Cybersecurity Awareness (TL&CA) ranged from 0.750 to 0.832, indicating robust indicator reliability.

**Table 2.** Measurement Model Quality Metrics.

| Construct                                                  | Loadings | Cronbach's $\alpha$ | AVE  | CR   |
|------------------------------------------------------------|----------|---------------------|------|------|
| Technological Literacy and Cybersecurity Awareness (TL&CA) |          | 0.86                | 0.61 | 0.89 |
| TL&CA1                                                     | 0.781    |                     |      |      |
| TL&CA2                                                     | 0.821    |                     |      |      |
| TL&CA3                                                     | 0.812    |                     |      |      |
| TL&CA4                                                     | 0.750    |                     |      |      |
| TL&CA5                                                     | 0.832    |                     |      |      |
| Social Media Engagement Intensity (SMEI)                   |          | 0.88                | 0.62 | 0.9  |
| SMEI1                                                      | 0.812    |                     |      |      |
| SMEI2                                                      | 0.851    |                     |      |      |
| SMEI3                                                      | 0.792    |                     |      |      |
| SMEI4                                                      | 0.842    |                     |      |      |
| SMEI5                                                      | 0.821    |                     |      |      |
| Digital Identity Visibility (DIV)                          |          | 0.79                | 0.57 | 0.85 |

|                                               |       |       |      |      |      |
|-----------------------------------------------|-------|-------|------|------|------|
|                                               | DIV1  | 0.831 |      |      |      |
|                                               | DIV2  | 0.783 |      |      |      |
|                                               | DIV3  | 0.811 |      |      |      |
| AI-Mediated Interactions (AIMI)               |       |       | 0.85 | 0.55 | 0.88 |
|                                               | AIMI1 | 0.823 |      |      |      |
|                                               | AIMI2 | 0.772 |      |      |      |
|                                               | AIMI3 | 0.823 |      |      |      |
|                                               | AIMI4 | 0.791 |      |      |      |
| Cultural Norms and Social Expectations (CNSE) |       |       | 0.87 | 0.58 | 0.9  |
|                                               | CNSE1 | 0.762 |      |      |      |
|                                               | CNSE2 | 0.813 |      |      |      |
|                                               | CNSE3 | 0.792 |      |      |      |
|                                               | CNSE4 | 0.831 |      |      |      |
|                                               | CNSE5 | 0.771 |      |      |      |
| Exposure to Digital Harassment (EDH)          |       |       | 0.9  | 0.64 | 0.92 |
|                                               | EDH1  | 0.822 |      |      |      |
|                                               | EDH2  | 0.853 |      |      |      |
|                                               | EDH3  | 0.842 |      |      |      |
|                                               | EDH4  | 0.782 |      |      |      |
|                                               | EDH5  | 0.843 |      |      |      |
| Mental Health Impact (MHI)                    |       |       | 0.85 | 0.55 | 0.88 |
|                                               | MHI1  | 0.771 |      |      |      |
|                                               | MHI2  | 0.822 |      |      |      |
|                                               | MHI3  | 0.843 |      |      |      |
|                                               | MHI4  | 0.792 |      |      |      |
| E-Learning Continuity (ELC)                   |       |       | 0.87 | 0.6  | 0.9  |
|                                               | ELC1  | 0.822 |      |      |      |
|                                               | ELC2  | 0.794 |      |      |      |
|                                               | ELC3  | 0.831 |      |      |      |
|                                               | ELC4  | 0.814 |      |      |      |

Similarly, items measuring Social Media Engagement Intensity (SMEI) showed loadings ranging from 0.792 to 0.851, indicating that these indicators consistently capture the intensity of students' social media use. Constructs with fewer items, such as Digital Identity Visibility (DIV), also demonstrated strong loadings (0.783–0.831), supporting the stability and clarity of these reflective measures.

In addition, internal consistency reliability was further supported by Cronbach's  $\alpha$  and Composite Reliability (CR), with all constructs exceeding the recommended minimum of 0.70. Notably, Exposure to Digital Harassment (EDH) displayed excellent reliability ( $\alpha = 0.90$ ; CR = 0.92), reflecting the consistency of students' reported experiences across cyberbullying, cyber hate, sexual extortion, and digital stalking.

Furthermore, comparable reliability levels were observed for Cultural Norms and Social Expectations (CNSE) ( $\alpha = 0.87$ ; CR = 0.90) and E-Learning Continuity (ELC) ( $\alpha = 0.87$ ; CR = 0.90). Even constructs with moderate item counts, such as Mental Health Impact (MHI) and AI-Mediated Interactions (AIMI), demonstrated acceptable reliability, with CR values of 0.88.

Likewise, convergent validity was supported by the Average Variance Extracted (AVE), with all constructs surpassing the minimum threshold of 0.50. AVE values ranged from 0.55 to 0.64, indicating that each construct explains more than half of the variance in its observed indicators.

Additionally, EDH showed the strongest convergent validity (AVE = 0.64). At the same time, both TL&CA and SMEI demonstrated strong convergent validity (AVEs of 0.61

and 0.62, respectively). Additionally, the combined evidence indicates strong reliability and validity of the measurement model, providing a robust foundation for subsequent evaluation of the structural model. Figure 1 presents the structural model with the estimated path coefficients, illustrating the relationships among the associated factors, EDH, and student outcomes. (See Figure 4)

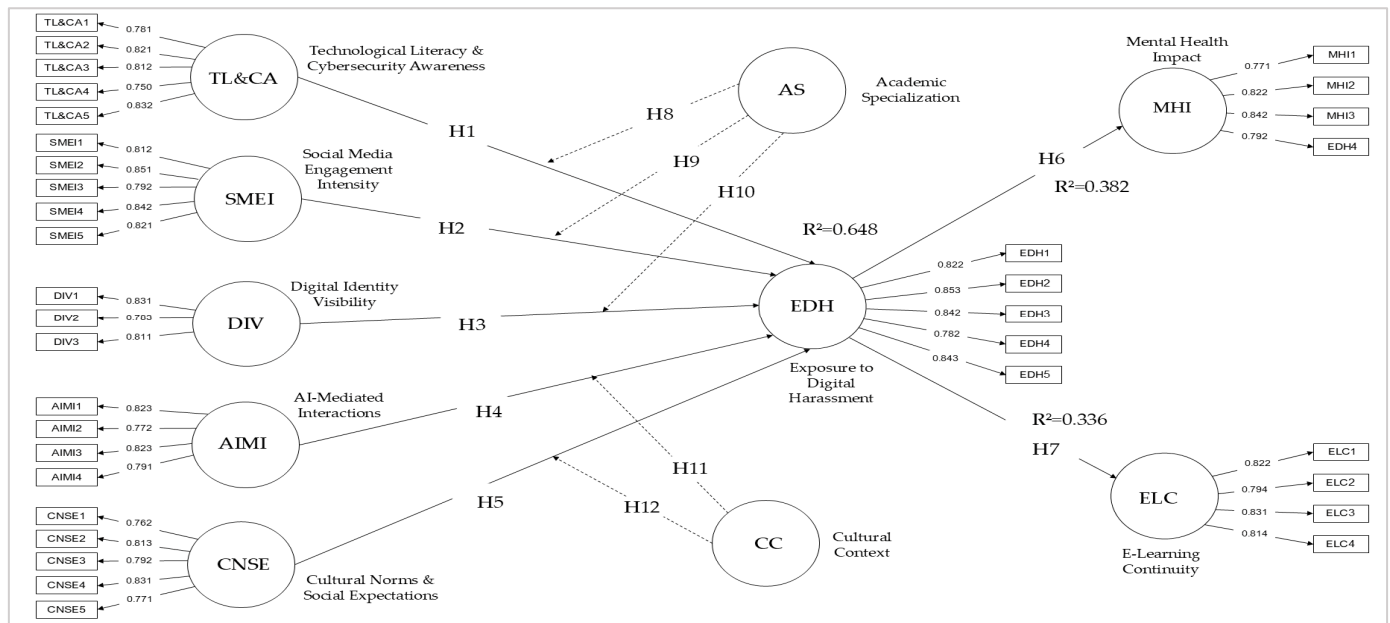


Figure 4. Statistical model.

Table 3 presents the assessment of discriminant validity using the Fornell–Larcker criterion, and Table 4 shows that all constructs exhibit satisfactory discriminant validity. For each latent variable, the square root of AVE, shown on the diagonal, is greater than the correlations with all other constructs in the corresponding row and column. This indicates that each construct accounts for more variance within its indicators than across indicators of other constructs—an essential requirement for reflective measurement models.

Table 3. Fornell–Larcker Criterion Matrix.

| Construct | TL&CA | SMEI  | DIV   | AIMI  | CNSE  | EDH   | MHI   | ELC   |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|
| TL&CA     | 0.781 |       |       |       |       |       |       |       |
| SMEI      | 0.42  | 0.787 |       |       |       |       |       |       |
| DIV       | 0.38  | 0.51  | 0.755 |       |       |       |       |       |
| AIMI      | 0.40  | 0.49  | 0.48  | 0.741 |       |       |       |       |
| CNSE      | 0.36  | 0.55  | 0.50  | 0.52  | 0.761 |       |       |       |
| EDH       | 0.44  | 0.62  | 0.54  | 0.58  | 0.60  | 0.800 |       |       |
| MHI       | 0.46  | 0.47  | 0.43  | 0.46  | 0.48  | 0.57  | 0.742 |       |
| ELC       | 0.41  | 0.45  | 0.40  | 0.42  | 0.44  | 0.50  | 0.52  | 0.775 |

For example, TL&CA exhibited a square root of AVE of 0.781, which exceeds its correlations with SMEI (0.42), DIV (0.38), AIMI (0.40), CNSE (0.36), EDH (0.44), MHI (0.46), and ELC (0.41); similar patterns were observed for other constructs. Additionally, the square root of AVE for SMEI was 0.787, noticeably higher than its correlations with DIV (0.51), AIMI (0.49), CNSE (0.55), EDH (0.62), MHI (0.47), and ELC (0.45). Likewise, DIV demonstrated a diagonal value of 0.755, exceeding its correlations with SMEI (0.51), AIMI (0.48), CNSE (0.50), EDH (0.54), MHI (0.43), and ELC (0.40).

Furthermore, the same pattern holds for AIMI, CNSE, and EDH. Notably, EDH had a diagonal loading of 0.800, which was higher than its correlations with all other constructs, including SMEI (0.62) and CNSE (0.60), the strongest correlations. Hence, this confirms that although these correlations are theoretically related, they remain empirically distinct.

In addition, the outcome variables—MHI and ELC—satisfied the Fornell–Larcker criterion, with diagonal loadings of 0.742 and 0.775, respectively, exceeding all interconstruct correlations. Collectively, these results provide strong evidence that all constructs in the model meet the discriminant validity requirements according to the Fornell–Larcker criterion. Similarly, the findings support the conclusion that each latent variable is unique and captures a distinct conceptual domain, thereby justifying its use in the subsequent structural model analysis.

**Table 4.** Heterotrait–Monotrait Ratio (HTMT) Matrix.

| Construct | TL&CA | SMEI | DIV  | AIMI | CNSE | EDH  | MHI  | ELC  |
|-----------|-------|------|------|------|------|------|------|------|
| TL&CA     |       |      |      |      |      | 0.58 |      |      |
| SMEI      |       |      | 0.63 | 0.60 | 0.66 | 0.74 |      |      |
| DIV       |       |      |      | 0.59 | 0.62 | 0.68 |      |      |
| AIMI      |       |      |      |      | 0.64 | 0.70 |      |      |
| CNSE      |       |      |      |      |      | 0.72 |      |      |
| EDH       |       |      |      |      |      |      | 0.68 | 0.62 |
| MHI       |       |      |      |      |      |      |      |      |
| ELC       |       |      |      |      |      |      |      |      |

As presented in Table 4, the HTMT values between the associated factor variables and exposure are moderate but acceptable. For instance, the HTMT value between SMEI and EDH is 0.74. In contrast, the values for DIV (0.68), AIMI (0.70), and CN&SE (0.72) also remain below the threshold of 0.85. These results are theoretically expected, as these associated factors are conceptually related to students’ digital behaviours and social experiences, yet they remain distinct constructs. Similarly, the HTMT value between EDH and MHI is 0.68, indicating a meaningful but not excessive association, consistent with theoretical expectations of the negative psychological consequences of EDH. The HTMT value between ELC remains within acceptable limits (0.62), indicating that although exposure influences ELC persistence, the two constructs measure distinct aspects of the student experience.

As presented in Table 5, items associated with TL&CA show substantially higher loadings on their own construct (0.75–0.83) than on any other construct, where cross-loadings range only between 0.26 and 0.38. This pattern is consistent across all TL&CA items, confirming that the indicators uniquely measure students’ technological capabilities and cybersecurity awareness.

**Table 5.** Indicator Loadings and Cross-Loadings.

| Item   | TL&CA | SMEI | DIV  | AIMI | CNSE | EDH  | MHI  | ELC  |
|--------|-------|------|------|------|------|------|------|------|
| TL&CA1 | 0.78  | 0.32 | 0.28 | 0.30 | 0.27 | 0.33 | 0.29 | 0.26 |
| TL&CA2 | 0.82  | 0.35 | 0.30 | 0.32 | 0.29 | 0.36 | 0.32 | 0.28 |
| TL&CA3 | 0.80  | 0.33 | 0.29 | 0.31 | 0.28 | 0.35 | 0.30 | 0.27 |
| TL&CA4 | 0.75  | 0.30 | 0.27 | 0.29 | 0.26 | 0.31 | 0.28 | 0.25 |
| TL&CA5 | 0.83  | 0.36 | 0.32 | 0.34 | 0.31 | 0.38 | 0.33 | 0.30 |
| SMEI1  | 0.33  | 0.81 | 0.45 | 0.43 | 0.47 | 0.53 | 0.39 | 0.37 |
| SMEI2  | 0.34  | 0.85 | 0.47 | 0.45 | 0.48 | 0.55 | 0.41 | 0.39 |
| SMEI3  | 0.31  | 0.79 | 0.43 | 0.41 | 0.46 | 0.51 | 0.38 | 0.34 |

|       |      |      |      |      |      |      |      |      |
|-------|------|------|------|------|------|------|------|------|
| SMEI4 | 0.32 | 0.84 | 0.46 | 0.44 | 0.47 | 0.54 | 0.40 | 0.38 |
| SMEI5 | 0.30 | 0.80 | 0.42 | 0.40 | 0.45 | 0.50 | 0.37 | 0.35 |
| DIV1  | 0.29 | 0.47 | 0.83 | 0.44 | 0.46 | 0.50 | 0.38 | 0.35 |
| DIV2  | 0.27 | 0.45 | 0.78 | 0.41 | 0.44 | 0.47 | 0.36 | 0.33 |
| DIV3  | 0.28 | 0.46 | 0.81 | 0.43 | 0.45 | 0.49 | 0.37 | 0.35 |
| AIMI1 | 0.30 | 0.43 | 0.44 | 0.80 | 0.47 | 0.52 | 0.36 | 0.34 |
| AIMI2 | 0.28 | 0.42 | 0.41 | 0.77 | 0.45 | 0.50 | 0.35 | 0.32 |
| AIMI3 | 0.32 | 0.44 | 0.43 | 0.82 | 0.48 | 0.54 | 0.39 | 0.36 |
| AIMI4 | 0.29 | 0.40 | 0.41 | 0.79 | 0.46 | 0.51 | 0.34 | 0.33 |
| CNSE1 | 0.27 | 0.48 | 0.46 | 0.47 | 0.76 | 0.57 | 0.39 | 0.36 |
| CNSE2 | 0.29 | 0.49 | 0.48 | 0.48 | 0.81 | 0.59 | 0.42 | 0.38 |
| CNSE3 | 0.28 | 0.47 | 0.46 | 0.46 | 0.79 | 0.58 | 0.41 | 0.37 |
| CNSE4 | 0.30 | 0.50 | 0.49 | 0.48 | 0.83 | 0.61 | 0.43 | 0.39 |
| CNSE5 | 0.26 | 0.46 | 0.44 | 0.45 | 0.77 | 0.55 | 0.38 | 0.35 |
| EDH1  | 0.33 | 0.53 | 0.50 | 0.54 | 0.56 | 0.82 | 0.48 | 0.44 |
| EDH2  | 0.34 | 0.54 | 0.51 | 0.55 | 0.58 | 0.85 | 0.50 | 0.46 |
| EDH3  | 0.32 | 0.52 | 0.49 | 0.53 | 0.55 | 0.80 | 0.47 | 0.43 |
| EDH4  | 0.30 | 0.50 | 0.47 | 0.51 | 0.53 | 0.78 | 0.45 | 0.41 |
| EDH5  | 0.36 | 0.56 | 0.53 | 0.57 | 0.60 | 0.84 | 0.49 | 0.45 |
| MHI1  | 0.33 | 0.41 | 0.38 | 0.39 | 0.40 | 0.48 | 0.77 | 0.43 |
| MHI2  | 0.36 | 0.43 | 0.40 | 0.41 | 0.42 | 0.50 | 0.82 | 0.46 |
| MHI3  | 0.31 | 0.39 | 0.37 | 0.38 | 0.39 | 0.46 | 0.80 | 0.44 |
| MHI4  | 0.30 | 0.38 | 0.36 | 0.37 | 0.38 | 0.45 | 0.79 | 0.42 |
| ELC1  | 0.28 | 0.40 | 0.37 | 0.38 | 0.39 | 0.44 | 0.45 | 0.82 |
| ELC2  | 0.30 | 0.38 | 0.35 | 0.36 | 0.37 | 0.43 | 0.44 | 0.79 |
| ELC3  | 0.29 | 0.39 | 0.36 | 0.37 | 0.38 | 0.45 | 0.46 | 0.83 |
| ELC4  | 0.31 | 0.40 | 0.38 | 0.39 | 0.40 | 0.47 | 0.47 | 0.81 |

Similarly, items measuring SMEI exhibit high loadings on their own construct (0.79–0.85) and noticeably weaker cross-loadings with other constructs, ranging from 0.30 to 0.55. These results indicate that SMEI items associationively capture students' engagement intensity with digital platforms without substantially overlapping with other constructs, such as DIV or AIMI. Likewise, items for DIV and AIMI follow the same pattern: primary loadings remain strong (DIV = 0.78–0.83; AIMI = 0.77–0.82), while cross-loadings remain considerably lower, reinforcing their conceptual independence within the model.

Additionally, the CNSE also demonstrates strong discriminant validity, with primary loadings ranging from 0.76 to 0.83 and cross-loadings that never exceed the highest loading for any item. This finding supports the argument that CNSE captures distinct cultural pressures, reporting expectations, and social constraints that shape experiences of digital harassment. In addition, items for EDH—which represent behaviours such as cyberbullying, cyber hate, sexual extortion, and digital stalking—exhibited high loadings (0.78–0.85) and consistently lower cross-loadings, indicating that the indicators collectively reflect a coherent construct relating to online harassment.

Furthermore, the outcome constructs—MHI and ELC—also demonstrate strong discriminant validity. MHI loadings fall between 0.77 and 0.82, while ELC loadings range from 0.79 to 0.83, with all cross-loadings remaining well below the primary loading values. This pattern indicates that the indicators for emotional strain, anxiety, and reduced learning persistence are uniquely associated with their respective constructs rather than overlapping with correlated variables.

As presented in Table 6, a structural model assessment indicates strong support for most hypothesised relationships within the proposed framework. Among the five associated factors of digital harassment, four demonstrated significant positive associations

with students' exposure to online harm. AIMI (H4) showed the strongest statistical association with exposure ( $\beta = 0.31$ ,  $p < 0.001$ ), suggesting that algorithmically driven communication, automated content, and AI-enhanced digital environments significantly increase the likelihood of encountering harmful behaviour.

**Table 6.** Hypotheses Evaluation.

| Assumed Hypotheses                                             | $\beta$ | T    | $p$ Values | Evaluation    |
|----------------------------------------------------------------|---------|------|------------|---------------|
| H1: TL&CA $\rightarrow$ EDH                                    | −0.21   | 4.32 | 0.000      | Supported     |
| H2: SMEI $\rightarrow$ EDH                                     | 0.28    | 6.10 | 0.000      | Supported     |
| H3: DIV $\rightarrow$ EDH                                      | 0.24    | 5.02 | 0.000      | Supported     |
| H4: AIMI $\rightarrow$ EDH                                     | 0.31    | 6.45 | 0.000      | Supported     |
| H5: CNSE $\rightarrow$ EDH                                     | 0.27    | 5.89 | 0.000      | Supported     |
| H6: EDH $\rightarrow$ MHI                                      | −0.33   | 7.21 | 0.000      | Supported     |
| H7: EDH $\rightarrow$ ELC                                      | −0.29   | 6.44 | 0.000      | Supported     |
| H8: AS $\times$ Technological Literacy $\rightarrow$ Exposure  | 0.07    | 1.92 | 0.055      | Not Supported |
| H9: AS $\times$ Social Media Engagement $\rightarrow$ Exposure | 0.11    | 2.45 | 0.014      | Supported     |
| H10: AS $\times$ DIV $\rightarrow$ Exposure                    | 0.09    | 2.11 | 0.035      | Supported     |
| H11: CC $\times$ AIMI $\rightarrow$ Exposure                   | 0.13    | 2.78 | 0.006      | Supported     |
| H12: CC $\times$ Cultural Norms $\rightarrow$ Exposure         | 0.05    | 1.48 | 0.139      | Not Supported |

Additionally, SMEI (H2) also exhibited a substantial positive association ( $\beta = 0.28$ ,  $p < 0.001$ ), highlighting the role of high-frequency online activity and social platform interactions in increasing vulnerability to harassment. Likewise, CNSE (H5) ( $\beta = 0.27$ ,  $p < 0.001$ ) and DIV (H3) ( $\beta = 0.24$ ,  $p < 0.001$ ) significantly increased exposure levels, indicating that cultural pressures, social constraints, and the openness of one's digital identity contribute meaningfully to being associated with higher levels of patterns. In contrast, TL&CA (H1) showed a significant negative relationship with EDH ( $\beta = -0.21$ ,  $p < 0.001$ ). This finding confirms that students with more substantial knowledge of cybersecurity practices, online privacy management, and safe digital behaviours are less likely to experience harassment.

Furthermore, the protective association of technological literacy underscores the importance of digital hygiene education and student training to reduce the associated risks. The consequences of exposure to digital harassment were also strongly supported. Exposure to digital harassment (EDH) significantly worsened mental health impact (MHI) (H6;  $\beta = -0.33$ ,  $p < 0.001$ ), indicating that higher levels of harassment are associated with increased emotional distress, anxiety, and poorer psychological well-being. Given the negative valence of the MHI scale, this negative coefficient reflects more severe adverse mental health outcomes as exposure is associated with higher levels.

Exposure also significantly reduced ELC (H7) ( $\beta = -0.29$ ,  $p < 0.001$ ), demonstrating that students who experience harassment are more likely to disengage from digital learning environments, skip online sessions, or withdraw from online activities. These results emphasise the severe academic and emotional implications of digital harassment within AI-mediated educational contexts.

Moreover, regarding moderation associations, Academic Specialisation (AS) partially moderates the relationships among the associated factors and exposure to harassment. While its interaction with technological literacy (H8) was not statistically significant ( $\beta = 0.07$ ,  $p = 0.055$ ), AS significantly moderated the associations of Social Media Engagement (H9) ( $\beta = 0.11$ ,  $p = 0.014$ ) and DIV (H10) ( $\beta = 0.09$ ,  $p = 0.035$ ). These findings suggest that students across fields of study may experience and interpret digital engagement and visibility in distinct ways, thereby influencing their vulnerability to online harm.

Cultural Context (CC) also showed selective moderation associations. It significantly strengthened the relationship between AIMI and Exposure (H11) ( $\beta = 0.13$ ,  $p = 0.006$ ),

indicating that cultural differences may shape how students experience AI-driven digital environments and their higher levels of association. However, CC did not significantly moderate the relationship between cultural norms and exposure (H12) ( $\beta = 0.05$ ,  $p = 0.139$ ), suggesting that the direct association of cultural norms is sufficiently strong and consistent across cultural groups, regardless of contextual variations.

Taken together, these results provide strong empirical support for the proposed model. Most associated factors significantly influence exposure to digital harassment; exposure, in turn, significantly predicts mental health outcomes and e-learning persistence. The moderating roles of AS and CC highlight significant socio-academic variations associated with higher levels of patterns, adding nuance to the understanding of how digital harassment operates within AI-mediated learning environments.

As presented in Table 7, the structural model quality metrics, including  $R^2$ ,  $f^2$ , and  $Q^2$  values, collectively assess the explanatory power, association sizes, and predictive relevance of the proposed PLS-SEM model. Overall, the findings indicate that the model demonstrates strong explanatory capacity and meaningful predictive accuracy for the central constructs of interest—EDH, MHI, and ELC.

**Table 7.** Structural Model Quality Metrics.

| Endogenous Construct | Correlate Variable   | $R^2$ | $f^2$ | $Q^2$ |
|----------------------|----------------------|-------|-------|-------|
| EDH (EDH)            | TL&CA (TL&CA)        | 0.648 | 0.06  | 0.411 |
|                      | SMEI (SMEI)          |       | 0.12  |       |
|                      | DIV (DIV)            |       | 0.09  |       |
|                      | AIMI (AIMI)          |       | 0.14  |       |
|                      | CNSE (CNSE)          |       | 0.11  |       |
|                      | AS (Moderator paths) |       | 0.02  |       |
|                      | CC (Moderator paths) |       | 0.03  |       |
| MHI (MHI)            | EDH (EDH)            | 0.382 | 0.18  | 0.272 |
| ELC (ELC)            | EDH (EDH)            | 0.336 | 0.16  | 0.234 |

Additionally, the model explains a substantial proportion of variance in EDH, with an  $R^2$  value of 0.648, indicating that approximately 65% of the variability in exposure is accounted for by the five associated factors and the moderating variables. This level of explanatory power is considered high for behavioural and educational research, especially within complex digital environments.

Furthermore, the association sizes ( $f^2$ ) indicate that AIMI exerted the most significant relationship with EDH ( $f^2 = 0.14$ ), followed by SMEI ( $f^2 = 0.12$ ) and CNSE ( $f^2 = 0.11$ ). Likewise, DIV ( $f^2 = 0.09$ ) demonstrated a meaningful association, whereas TL&CA (TL&CA) exhibited a smaller yet significant protective association ( $f^2 = 0.06$ ). The moderating variables—AS ( $f^2 = 0.02$ ) and CC ( $f^2 = 0.03$ )—showed small but theoretically relevant contributions, consistent with prior research on contextual and disciplinary influences in digital behaviour.

Predictive relevance, as assessed by  $Q^2$ , further validates the model. Moreover, the  $Q^2$  value for EDH ( $Q^2 = 0.411$ ) is well above zero, indicating strong predictive relevance and confirming that the model accurately predicts the levels of digital harassment exposure among students. This finding underscores the robustness of the structural pathways linking behavioural, technological, and cultural associated factors with harassment outcomes.

The outcome constructs—MHI (MHI) and ELC (ELC)—also demonstrated meaningful explanatory power. EDH explained 38.2% of the variance in mental health outcomes ( $R^2 = 0.382$ ) and 33.6% of the variance in ELC ( $R^2 = 0.336$ ). These results indicate moderate to strong explanatory power, highlighting the central role of harassment experiences in

shaping both psychological well-being and students' ability to sustain participation in online learning.

Furthermore, association sizes confirm these relationships:  $f^2 = 0.18$  for MHI and  $f^2 = 0.16$  for ELC, both indicating moderate associations. Predictive relevance values ( $Q^2 = 0.272$  for MHI and  $Q^2 = 0.234$  for ELC) support the conclusion that EDH meaningfully predicts declines in mental health and disruptions in online learning engagement. Hence, the structural model demonstrates strong explanatory capabilities, meaningful association sizes, and robust predictive relevance. The results validate the theoretical framework proposed in this study and highlight the central role of exposure to digital harassment in shaping students' psychological and academic experiences within AI-mediated educational environments.

## 5. Discussion

This study examined the behavioural, technological, and cultural factors associated with university students' EDH in AI-mediated learning environments. The findings extend prior research by demonstrating how AI-driven interactions, digital identity patterns, and social engagement behaviours collectively contribute to online harm. This is consistent with earlier work on technology-mediated education (e.g., [20,66,67]) and contributes to current debates on how digital societies and algorithmic systems reshape social reality. The results confirm that digital harassment remains a pervasive challenge for students, particularly in developing countries undergoing rapid digital expansion, where platform penetration and AI-supported tools grow faster than regulatory, pedagogical, and ethical safeguards.

The strong reliability and validity of the measurement model further confirm that constructs such as technological literacy, cultural norms, identity visibility, and harassment exposure are conceptually distinct and empirically robust, consistent with prior psychometric research in digital safety and online, which is associated with higher levels of these domains (e.g., [28,29,37,53]). Moreover, a key empirical contribution of this study is the identification of AI-Mediated Interactions (AIMI) as the strongest correlation of EDH. This finding aligns with growing concerns that algorithmic systems may unintentionally amplify harmful content, facilitate repeated or unsolicited contact, or weaken traditional cues of accountability through automated mediation (e.g., [30,54,68]). In line with prior research on algorithmic curation, AI-driven recommendation and ranking mechanisms can increase exposure to extreme or undesirable interactions, thereby intensifying associations with factors beyond individual behavioural factors.

The significant association between Social Media Engagement Intensity (SMEI) and frequent participation in online spaces is consistent with earlier findings showing that frequent participation in online spaces is associated with greater exposure to cyber aggression [3,5,21]. Additionally, High posting activity, real-time interaction, and presence across multiple platforms enhance visibility and accessibility to potential perpetrators. In AI-mediated environments, this visibility is further shaped by algorithmic amplification, which surfaces, promotes, or recommends highly active users. Consequently, SMEI is associated with higher levels of mechanisms operating at the intersection of user behaviour and platform-driven exposure dynamics.

Similarly, Digital Identity Visibility (DIV) was found to significantly increase vulnerability to harassment. This result is consistent with studies demonstrating that publicly accessible profiles, extensive self-disclosure, and weak privacy controls are associated with higher levels of stalking, impersonation, and targeted harassment [28,30,33]. Even limited publicly available information—such as identifiable images or activity traces—can be exploited by malicious actors. The findings confirm that visible and weakly protected

digital identities constitute a persistent structural vulnerability within AI-enhanced educational ecosystems.

The influence of Cultural Norms and Social Expectations (CNSE) highlights the central role of stigma, honour-related concerns, and reporting cultures in shaping prolonged exposure to digital harassment. In many contexts, students may refrain from reporting harassment due to fear of social repercussions, reputational damage, or perceptions of institutional inaction—patterns well documented in prior research [3,5,16,20,55,66]. These socio-cultural pressures can normalise online aggression and discourage help-seeking, allowing harmful behaviours to persist and escalate. The present findings underscore that digital harassment cannot be understood solely as a technological or individual issue but must be situated within broader cultural and social frameworks that shape tolerance, silence, and vulnerability.

The protective association of Technological Literacy and Cybersecurity Awareness (TL&CA) parallels findings from digital resilience research, which consistently show that knowledge of privacy controls, threat detection, and secure communication reduces victimisation [15,53]. Students equipped with digital safety skills are better able to navigate AI-mediated environments, manage exposure, and respond to emerging threats. Within developing contexts, TL&CA also reflects a critical dimension of digital inequality: disparities in skills and awareness translate into unequal protection against algorithmically amplified associations.

The study further confirms that EDH has significant adverse associations with mental health, consistent with extensive evidence linking online harassment to anxiety, emotional exhaustion, fear, and depressive symptoms [16,17,55]. In addition, the negative association with e-learning continuity aligns with research showing that harassment leads to disengagement, avoidance of online platforms, and reduced academic motivation [6,15]. In AI-enhanced learning environments—where digital participation is central to assessment, collaboration, and communication—harassment can reshape educational trajectories and exacerbate inequalities in academic persistence and success. Moderation analyses provide additional insight into contextual variation. Academic specialisation moderated the associations of social media engagement and digital identity visibility, suggesting that discipline-specific cultures and patterns of digital exposure shape harassment [69].

By contrast, the absence of moderation by specialisation for TL&CA indicates that cybersecurity awareness offers broadly protective benefits across fields. Cultural context moderated the relationship between AI-mediated interactions and EDH, reinforcing evidence that perceptions of AI, platform trust, and digital norms vary across cultural environments. However, cultural context did not moderate the association of CNSE, indicating that stigma and reporting pressures may operate as shared structural constraints rather than context-specific cultural traits, consistent with cross-cultural cyberbullying research [70–72].

Within developing countries, the results suggest that students often engage with AI-enhanced platforms through shared digital infrastructure, limited privacy controls, and fragmented governance frameworks. These conditions can increase both the visibility and persistence of harmful online interactions. High levels of social media engagement and digital identity visibility further increase the association, particularly when online participation substitutes for academic and social interaction. In such contexts, students are frequently exposed to large digital audiences in environments characterised by weak cyber-regulation, a pattern consistent with earlier studies [5,9].

Cultural norms surrounding reporting, combined with higher tolerance of online aggression, may further intensify vulnerability. Social stigma, honour-related concerns, and limited trust in institutional response mechanisms can discourage disclosure, allowing

harassment to persist and escalate. The absence of a moderating association of cultural context on cultural norms suggests that these pressures reflect shared structural constraints rather than isolated cultural characteristics. By contrast, technological literacy and cybersecurity awareness emerged as protective factors, indicating that individual digital competence can mitigate the effects associated with limited institutional safeguards. In developing contexts, gaps in digital skills therefore represent a form of digital inequality that translates into unequal exposure to algorithmically amplified harm [45,71].

Overall, these findings indicate that digital harassment in AI-driven learning environments in developing countries should be understood as a structural sociotechnical issue. It is shaped by rapid AI adoption, constrained governance, cultural barriers to reporting, and unequal access to knowledge about digital safety, rather than by isolated acts of individual misconduct.

## 6. Theoretical and Practical Contributions

This study advances the literature on digital harassment and educational technology by extending established theories of online victimisation into AI-mediated learning environments. Consistent with prior research, the findings confirm that social media engagement intensity, digital identity visibility, and cultural norms remain core factors associated with exposure to harassment. This continuity demonstrates that fundamental behavioural and socio-cultural mechanisms identified in social media contexts remain operative in higher education, even as learning increasingly shifts to AI-enhanced platforms [3,4,6,24,30]. The study therefore reinforces theoretical links between digital interaction intensity, visibility, and vulnerability, while situating them within compulsory academic participation settings.

A key theoretical contribution lies in conceptualising AI-mediated interactions not as a new form of harassment, but as a layer that amplifies and reshapes exposure conditions. Algorithmic mechanisms—such as content recommendation, ranking, automated prompts, and visibility optimisation—reconfigure how often, how long, and with whom students interact online. Hence, AI alters the scale, persistence, and reach of harmful encounters without changing the underlying social behaviours that constitute harassment. This distinction refines existing models by separating harassment acts from the algorithmic infrastructures that intensify their impact.

Moreover, the study empirically demonstrates how algorithmic affordances intensify established processes of victimisation. The strong statistical association with AI-Mediated Interactions (AIMI) shows that algorithmic curation can increase repeated contact with harmful actors, prolong the circulation of abusive content, and reduce user control over digital visibility. These findings extend sociotechnical and routine activity perspectives, which emphasise exposure, accessibility, and visibility as key factors associated with higher levels of conditions, by incorporating algorithmic opacity and diminished user agency as central theoretical elements [15,33,38].

Notably, the study avoids framing AI as an autonomous statistical agent. Instead, AI-mediated interactions are theorised as contextual conditions that modify exposure dynamics. By modelling AI-mediated interactions alongside behavioural and cultural associated factors, the study demonstrates that algorithmic mediation adds explanatory value beyond traditional harassment models. This integrated framework strengthens theoretical precision and situates AI-related issues within the established empirical and conceptual foundations of digital victimisation research.

## 7. Conclusions

This study examined behavioural, technological, and cultural factors associated with university students' exposure to digital harassment (EDH) in AI-mediated learning environments using a large cross-sectional sample. The results indicate that higher social media engagement intensity, greater visibility of digital identity, stronger constraining cultural norms, and more frequent AI-mediated interactions are associated with higher reported exposure. In contrast, higher technological literacy and cybersecurity awareness are associated with lower exposure. These findings should be interpreted as correlational: they do not demonstrate that AI systems cause harassment or that harassment directly causes mental health decline or reduced e-learning continuity.

Rather, the evidence supports viewing EDH as a sociotechnical system characterised by higher levels of conditions, shaped by student behaviours, cultural constraints, and platform-mediated interaction structures. AI-mediated interactions showed the strongest association with exposure, suggesting that algorithmic curation and automated recommendations may influence exposure conditions (e.g., visibility, reach, persistence) within which harmful encounters can occur.

Moderation results also suggest context dependence by academic specialisation and cultural context, though some interaction associations were small or non-significant, warranting cautious generalisation. Key uncertainties remain due to cross-sectional self-report data and the use of a unified EDH construct that does not separate harassment subtypes. Next steps include longitudinal/panel designs, platform-level behavioural measures (logs, recommendation pathways, moderation records), and analyses of harassment subtypes to strengthen statistical inference and inform targeted interventions.

## 8. Limitations and Future Research

This study employed a cross-sectional design, although data were collected across multiple academic semesters. While robustness checks indicated temporal stability in the core constructs, this design limits the ability to capture changes in exposure to digital harassment over time. Given the rapid evolution of AI platforms, governance frameworks, and digital practices, future research should adopt longitudinal or panel-based designs to examine how shifts in algorithmic systems, platform policies, and user behaviours influence exposure to harassment in higher education.

Second, the study relied on self-reported data, which may be subject to recall bias, social desirability associations, or underreporting due to stigma surrounding harassment experiences. Future studies could enhance measurement validity by incorporating complementary data sources, such as digital trace data, platform behavioural logs, or institutional incident reports, to triangulate self-reported exposure and outcomes.

Although the sample included students from 33 developing and emerging countries, the analysis was conducted within a shared host-country educational context. While this approach ensured institutional and technological consistency, it limited the ability to capture finer-grained cross-national variation in regulatory environments, cultural norms, and reporting practices. Future research employing country-level, regional, or multi-host comparative designs could provide more nuanced insights into how local sociotechnical contexts shape the dynamics of digital harassment.

The study focused on the association between mental health and e-learning continuity, with mental health as the primary outcome. Future research should examine additional academic, social, and behavioural consequences, including academic performance, student retention, identity development, and institutional trust. Finally, modelling exposure as a unified construct may obscure differences among harassment subtypes. Future studies should adopt multidimensional, mixed-method approaches to clarify how AI-

mediated interactions shape specific exposure pathways and sociotechnical harm in AI-enhanced education.

**Funding:** This work was funded by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Project No. KFU260765].

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board at King Faisal University; approval code KFU-2025-ETHICS3829; approval date 2024-01-21.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The dataset is available at the following link: <https://2u.pw/mRgWI> (accessed on 18 January 2026).

**Conflicts of Interest:** The author declares no conflicts of interest.

## Appendix A

**Table A1.** Sample demographic characteristics.

| Cultural Context (CC) | Nationality  | Male |      | Female |      | Sum |       |
|-----------------------|--------------|------|------|--------|------|-----|-------|
|                       |              | n    | %    | n      | %    | n   | %     |
| Middle East           | Saudi Arabia | 139  | 6.36 | 167    | 7.64 | 306 | 14    |
|                       | Syria        | 141  | 6.45 | 163    | 7.46 | 304 | 13.91 |
|                       | Yemen        | 144  | 6.59 | 159    | 7.28 | 303 | 13.87 |
|                       | Egypt        | 127  | 5.81 | 173    | 7.92 | 300 | 13.73 |
|                       | Sudan        | 122  | 5.58 | 137    | 6.27 | 259 | 11.85 |
|                       | Palestine    | 98   | 4.49 | 101    | 4.62 | 199 | 9.11  |
|                       | Jordan       | 63   | 2.88 | 82     | 3.75 | 145 | 6.64  |
|                       | Bahrain      | 21   | 0.96 | 43     | 1.97 | 64  | 2.93  |
|                       | Tunisia      | 21   | 0.96 | 33     | 1.51 | 54  | 2.47  |
|                       | Kuwait       | 21   | 0.96 | 23     | 1.05 | 44  | 2.01  |
|                       | Oman         | 12   | 0.55 | 8      | 0.37 | 20  | 0.92  |
|                       | Qatar        | 5    | 0.23 | 12     | 0.55 | 17  | 0.78  |
|                       | Lebanon      | 1    | 0.05 | 15     | 0.69 | 16  | 0.73  |
|                       | Algeria      | 3    | 0.14 | 7      | 0.32 | 10  | 0.46  |
|                       | Morocco      | 2    | 0.09 | 1      | 0.05 | 3   | 0.14  |
| Eastern Asia          | Pakistan     | 12   | 0.55 | 18     | 0.82 | 30  | 1.37  |
|                       | India        | 13   | 0.59 | 15     | 0.69 | 28  | 1.28  |
|                       | Bangladesh   | 2    | 0.09 | 3      | 0.14 | 5   | 0.23  |
| Africa                | Uganda       | 10   | 0.46 | 1      | 0.05 | 11  | 0.5   |
|                       | Congo        | 7    | 0.32 | 3      | 0.14 | 10  | 0.46  |
|                       | Chad         | 6    | 0.27 | 2      | 0.09 | 8   | 0.37  |
|                       | Senegal      | 2    | 0.09 | 5      | 0.23 | 7   | 0.32  |
|                       | Cameroon     | 3    | 0.14 | 2      | 0.09 | 5   | 0.23  |
|                       | Niger        | 3    | 0.14 | 2      | 0.09 | 5   | 0.23  |
|                       | Benin        | 1    | 0.05 | 4      | 0.18 | 5   | 0.23  |
|                       | Burkina Faso | 2    | 0.09 | 3      | 0.14 | 5   | 0.23  |
|                       | Togo         | 3    | 0.14 | 1      | 0.05 | 4   | 0.18  |
|                       | Ghana        | 2    | 0.09 | 2      | 0.09 | 4   | 0.18  |
|                       | Guinea       | 1    | 0.05 | 3      | 0.14 | 4   | 0.18  |
|                       | Gambia       | 1    | 0.05 | 2      | 0.09 | 3   | 0.14  |
|                       | Kenya        | 1    | 0.05 | 2      | 0.09 | 3   | 0.14  |

|       |         |     |        |      |        |      |      |
|-------|---------|-----|--------|------|--------|------|------|
|       | Mali    | 1   | 0.05   | 1    | 0.05   | 2    | 0.09 |
|       | Nigeria | 1   | 0.05   | 1    | 0.05   | 2    | 0.09 |
| Total | 33      | 991 | 45.35% | 1194 | 54.65% | 2185 | 100% |

**Table A2.** Survey Items.

| Code                                                          | Item                                                                                                                                                                             |
|---------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1) Technological Literacy and Cybersecurity Awareness (TL&CA) |                                                                                                                                                                                  |
| TL&CA1                                                        | I know how to adjust privacy and security settings on the platforms I use for learning and communication.                                                                        |
| TL&CA2                                                        | I can recognise suspicious links, messages, or requests that may be phishing attempts or scams.                                                                                  |
| TL&CA3                                                        | I use strong, unique passwords and/or a password manager for my important academic or social accounts.                                                                           |
| TL&CA4                                                        | I regularly take steps to reduce my online activity, which is associated with higher stress levels, such as enabling two-factor authentication and updating my apps and devices. |
| TL&CA5                                                        | I understand how to protect my personal data when using public Wi-Fi networks or shared devices.                                                                                 |
| 2) Social Media Engagement Intensity (SMEI)                   |                                                                                                                                                                                  |
| SMEI1                                                         | I check social media immediately upon waking.                                                                                                                                    |
| SMEI2                                                         | I use social media within 15 min before going to sleep.                                                                                                                          |
| SMEI3                                                         | I use social media while having meals.                                                                                                                                           |
| SMEI4                                                         | I frequently post, comment, or send messages on social media during a typical day.                                                                                               |
| SMEI5                                                         | I frequently switch between multiple social media platforms within the same day.                                                                                                 |
| 3) Digital Identity Visibility (DIV)                          |                                                                                                                                                                                  |
| DIV1                                                          | My social or learning-related online profiles contain enough information for others to easily identify me (e.g., name, photo, ...).                                              |
| DIV2                                                          | A large portion of my online activity (posts, comments, or profile details) is visible to people beyond my close friends or classmates.                                          |
| DIV3                                                          | I often share personal information or opinions online that could attract attention from people I do not know well.                                                               |
| 4) AI-Mediated Interactions (AIMI)                            |                                                                                                                                                                                  |
| AIMI1                                                         | The platforms I use (e.g., social media, learning management systems, or apps) rely on algorithms to recommend content or accounts, shaping what I interact with.                |
| AIMI2                                                         | Automated ranking or recommendation systems influence which posts, messages, or people I notice and respond to online.                                                           |
| AIMI3                                                         | I use AI tools or chatbots (e.g., for explanations, drafting, or feedback) as part of my learning interactions.                                                                  |
| AIMI4                                                         | I frequently receive automated suggestions (e.g., people to follow, groups to join, or recommended peers or resources) that influence my online connections.                     |
| 5) Cultural Norms and Social Expectations (CNSE)              |                                                                                                                                                                                  |
| CNSE1                                                         | In my community, people often avoid reporting online harassment to prevent shame or negative social consequences.                                                                |
| CNSE2                                                         | Victims of online harassment are sometimes blamed for what happened, such as being too visible online.                                                                           |
| CNSE3                                                         | It is commonly expected that students should handle online harassment privately rather than seek institutional help.                                                             |
| CNSE4                                                         | Social norms in my context discourage speaking openly about harassment, especially when it involves sexual or reputational issues.                                               |
| CNSE5                                                         | Online aggression is often treated as normal or not taken seriously within my peer or community environment.                                                                     |
| 6) Exposure to Digital Harassment (EDH)                       |                                                                                                                                                                                  |
| EDH1                                                          | I have received insulting, humiliating, or threatening messages online.                                                                                                          |
| EDH2                                                          | I have been targeted by repeated unwanted contact or harassment, even after asking for it to stop.                                                                               |
| EDH3                                                          | I have experienced hateful or abusive comments targeting my identity (e.g., nationality, religion, ethnicity, or gender).                                                        |
| EDH4                                                          | I have experienced sexualised online harassment, sextortion pressure, or threats involving private or intimate content.                                                          |

|                                |                                                                                                                                         |
|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| EDH5                           | I have experienced digital stalking or monitoring, such as someone tracking, impersonating, or repeatedly observing my online accounts. |
| 7) Mental Health Impact (MHI)  |                                                                                                                                         |
| MHI1                           | Experiences of online harassment have made me feel anxious or on edge.                                                                  |
| MHI2                           | Online harassment has reduced my sense of safety when using digital platforms.                                                          |
| MHI3                           | After experiencing online harassment, I have felt distressed, upset, or emotionally exhausted.                                          |
| MHI4                           | Online harassment has negatively affected my concentration and daily functioning, including my academic focus.                          |
| 8) E-Learning Continuity (ELC) |                                                                                                                                         |
| ELC1                           | I intend to continue using online learning platforms for my courses whenever they are required or available.                            |
| ELC2                           | Even when online learning is challenging, I try to remain active by attending sessions, submitting assignments, and participating.      |
| ELC3                           | I feel able to sustain my engagement in online classes (e.g., forums, virtual sessions, or LMS activities) over time.                   |
| ELC4                           | Negative online experiences make me consider reducing my participation in online learning activities.                                   |

## References

1. Karkin, N.; Degirmenci, N.; Gülel, F.E. Gender-Based Psychological Side Associations of Online Education in the Pandemic Under the Lens of Administrative Burden. In *Handbook of Research on Exploring Gender Equity, Diversity, and Inclusion Through an Intersectional Lens*; IGI Global: Hershey, PA, USA, 2023.
2. Alotaibi, N.S. The impact of AI and LMS integration on the future of higher education: Opportunities, challenges, and strategies for transformation. *Sustainability* **2024**, *16*, 10357.
3. Ahmed, R. Cyber Harassment in the Digital Age: Trends, Challenges, and Countermeasures. *Radinka J. Sci. Syst. Lit. Rev.* **2024**, *2*, 442–450.
4. Alexander, S. Issues and Ideas, Deepfake Cyberbullying: The Psychological Toll on Students and Institutional Challenges of AI-Driven Harassment. *Clear. House A J. Educ. Strateg. Issues Ideas* **2025**, *98*, 36–50.
5. Ahsan, M.; Ali, F. The cost of speaking out: Cyber harassment and abuse against feminist activists in Pakistan. *Media Cult. Soc.* **2025**, *48*, 55–72.
6. Brax, D. *Victimisation in Swedish Higher Education Institutions: A Study of Harassment, Threats and Violence Against Researchers and Teachers*; Swedish Secretariat for Gender Research: Gothenburg, Sweden, 2024.
7. Omokhabi, U.S.; Omilani, M.A.; Adegbite, O.O.; Omokhabi, A.A. THE DARK SIDE OF ONLINE LEARNING: EXAMINING THE RISE OF CYBERBULLYING ON LEARNING INTEREST AND ENGAGEMENT OF STUDENTS IN THE AGE OF VIRTUAL CLASSROOMS. *UNIZIK J. Educ. Res. Policy Stud.* **2025**, *19*, 388–404.
8. Korneeva, E.; Strielkowski, W.; Krayneva, R.; Sherstobitova, A. Social health and psychological safety of students involved in online education during the covid-19 pandemic. *Int. J. Environ. Res. Public Health* **2022**, *19*, 13928.
9. Akeusola, B.N. Exploring the Impact of Technology-Facilitated Sexual Violence on the Well-being of Undergraduate Students in Ni-gerian Tertiary Institutions. *Indones. J. Soc. Sci. Res.* **2023**, *4*, 171–181.
10. Baeva, I.A.; Bordovskaia, N.V. The psychological safety of the educational environment and the psychological well-being of Russian secondary school pupils and teachers. *Psychol. Russ. State Art* **2015**, *8*, 86–99.
11. Myers, C.-A.; Cowie, H. Cyberbullying across the lifespan of education: Issues and interventions from school to university. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1217.
12. Majumder, A. Rise and impact of AI agents in the digital landscape. *Am. J. Intell. Syst.* **2025**, *14*, 10–17.
13. Oksanen, A.; Celuch, M.; Latikka, R.; Oksa, R.; Savela, N. Hate and harassment in academia: The rising concern of the online environment. *High. Educ.* **2022**, *84*, 541–567.
14. Kaur, A.; Maheshwari, S.; Varma, A. The digital escape: Examining the impact of cyberloafing on gossip-induced emotional exhaustion and the mediating role of self-esteem. In *Evidence-Based HRM: A Global Forum for Empirical Scholarship*; Emerald Publishing Limited: Leeds, UK, 2025.
15. Canales, A.; Webb, A. The how's and how often's of school harassment: How do they influence the academic achievement of adolescents over time? *Stud. Educ. Eval.* **2024**, *82*, 101358.

16. Swain, J. Exploring the Use of Digital Technologies and Stakeholder Perspectives of Online Harms and Harassment in the Higher Education Sector. Ph.D. Thesis, University of Plymouth, Plymouth, UK, 2024.
17. Waechter, N.; Meschik, M. Peer socialization of male adolescents in digital games: Achievement, competition, and harassment. *Communications* **2023**, *48*, 457–481.
18. Khattak, R. Widening Inclusion in Canadian higher education institutions towards a useful national data set for reporting on the Sustainable Development Goals. Ph.D. Thesis, University of Liverpool, Liverpool, UK, 2025.
19. Mutambara, P.D.E. UN Sustainable Development Goals (SDGs). In *Deploying Artificial Intelligence to Achieve the UN Sustainable Development Goals: Enablers, Drivers and Strategic Framework*; Springer: Cham, Switzerland, 2025; pp. 47–80.
20. Shaikh, F.B.; Meghji, A.F.; Dahri, N.A.; Kalhor, M.; Ayyasamy, R.K.; Subramaniam, S. Predicting cyber-harassment in the digital age: An integrated model of psychological, technological, and environmental factors using SEM. *Iran J. Comput. Sci.* **2025**, *8*, 2815–2840.
21. Gallegos, A.; Ampudia, L.G.; Córdova, H.M.; Londoño-Celis, W.; Mendoza, O.A.V.; Valencia, J. Cyberbullying in students: Forms of aggression, is associated with higher levels of factors, and educational responses in digital environments. *F1000Research* **2025**, *14*, 880.
22. Eslen-Ziya, H.; Giorgi, A.; Ahi, C.J. Digital vulnerabilities and online harassment of academics, consequences, and coping strategies: An exploratory analysis. *Fem. Media Stud.* **2024**, *24*, 1422–1427.
23. Ramkissoon, L. AI: Powering sustainable innovation in higher ed. In *The Evolution of Artificial Intelligence in Higher Education*; Emerald Publishing Limited: Leeds, UK, 2024; pp. 203–229.
24. Irene, B.N.; Irene, J.; Lockyer, J.; Dewitt, S. Degenerative ‘Affordance’ of Social Media in Family Business. *Systems* **2025**, *13*, 629.
25. Monteiro-Krebs, L. Recommendations in Academic Social Media: The Shaping of Scholarly Communication Through Algorithmic Mediation. Ph.D. Thesis, Federal University of Rio Grande Do Sul, Porto Alegre, Brazil, 2022.
26. Rachmad, Y.E. *Personalized Digital Influence Theory*; United Nations Economic and Social Council: New York, NY, USA, 2025.
27. Naseem, M.J.; Bardhan, M.R.; Singh, M.P. *The Digital Paradox of Childhood: Balancing Freedom & Protection in the Age of AI and Cybersecurity*; BlueRose Publishers Pvt Ltd: New Delhi, India, 2025.
28. Elrayah, M.; Jamil, S. Impact of digital literacy and online privacy concerns on cybersecurity behaviour: The moderating role of cybersecurity awareness. *Int. J. Cyber Criminol.* **2023**, *17*, 166–187.
29. Stegeman, H.M.; Are, C.; Poell, T. Strategic invisibility: How creators manage the is associated with higher levels of and constraints of online hyper (in) visibility. *Soc. Media + Soc.* **2024**, *10*, 20563051241244674.
30. Bhat, M. Toward an Ethic of Synthetic Relationality: Identity, Intimacy, and Is associated with higher levels of in AI-Mediated Roleplay Environments. In *Proceedings of the Eighth AAAI/ACM Conference on AI, Ethics, and Society (AIES-25)*; The AAAI Press: Washington, DC, USA, 2025.
31. Rosenberg, H.; Blongheim, M.; Porat, C.S.-B. Who in the World Is Generation Z? The Rise of Mobile Natives and Their Socio-Technological Identity. *Societies* **2025**, *15*, 314.
32. Huang, P.Y.; Deng, J.; Yang, Y.; Williams, S. Empowering Creators in the Fight Against Online Hate: A Qualitative Exploration of AI-Mediated Counterspeech Tools. In *Proceedings of the ACM on Human-Computer Interaction*; Association for Computing Machinery: New York, NY, USA, 2025; Volume 9, pp. 1–27.
33. Dahlström, A.; Dahlqvist, H.; Gådin, K.G. Co-occurring cyber and in-person victimisation of bullying and sexual harassment: The associations to depressive symptoms in Swedish adolescents. *BMC Public Health* **2025**, *25*, 786.
34. Elshaer, I.A.; Alnajdi, S.M.; Salem, M.A. AI-Assistive Technology Adoption and Mental Health Disorders in Visually Impaired University Students. *Electronics* **2025**, *14*, 4036.
35. Salem, M.A.; Sobaih, A.E.E. ADIDAS: An Examined Approach for Enhancing Cognitive Load and Attitudes towards Synchronous Digital Learning Amid and Post COVID-19 Pandemic. *Int. J. Environ. Res. Public Health* **2022**, *19*, 16972.
36. Salem, M.A.; Sobaih, A.E.E. A Quadruple “E” Approach for Associationive Cyber-Hygiene Behaviour and Attitude toward Online Learning among Higher-Education Students in Saudi Arabia amid COVID-19 Pandemic. *Electronics* **2023**, *12*, 2268.
37. Huang, B. Navigating digital divide: Exploring the influence of ideological and political education on cyber security and digital literacy amid information warfare. *Curr. Psychol.* **2024**, *43*, 23815–23836.
38. Gurudin, K.; Awasthi, P. Unveiling the Dark Side of the Internet: Cyber Harassment among Children and Its Consequences. *IAHRW Int. J. Soc. Sci. Rev.* **2025**, *13*, 81.
39. Wang, Q.; Camacho, I.; Jing, S.; Goel, A.K. Understanding the design space of AI-mediated social interaction in online learning: Challenges and opportunities. In *Proceedings of the ACM on Human-Computer Interaction*; Association for Computing Machinery: New York, NY, USA, 2022; Volume 6, pp. 1–26.

40. Goyal, N.; Park, L.; Vasserman, L. "You have to prove the threat is real": Understanding the needs of Female Journalists and Activists to Document and Report Online Harassment. In *CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*; Association for Computing Machinery: New York, NY, USA, 2022.
41. Idriss, M.M.; Chowdhury, R. 'Honour'-Based Stalking, in *Young People, Stalking Awareness and Domestic Abuse*; Springer: Cham, Switzerland, 2023; pp. 161–182.
42. Perreault, G.P. Digital hate: Normalization in management of online hostility. In *The Routledge Companion to Digital Journalism Studies*; Routledge: London, UK, 2023; pp. 406–416.
43. Johansen, K.B.H.; Tjørnhøj-Thomsen, T. 'It's like you're almost being exposed to a flasher'—Young people's experiences with gendered vulnerability and digital sexual harassment in Denmark. *Women's Stud. Int. Forum* **2024**, *105*, 102954.
44. James, T.L.; Zhang, J.; Li, H.; Ziegelmayer, J.L.; Villacis-Calderon, E.D. The moderating association of technology overload on the ability of online learning to meet students' basic psychological needs. *Inf. Technol. People* **2022**, *35*, 1364–1382.
45. Chen, G.M.; Pain, P.; Chen, V.Y.; Mekelburg, M.; Springer, N.; Troger, F. 'You really have to have a thick skin': A cross-cultural perspective on how online harassment influences female journalists. *Journalism* **2020**, *21*, 877–895.
46. Powell, A.; Scott, A.J.; Henry, N. Digital harassment and abuse: Experiences of sexuality and gender minority adults. *Eur. J. Criminol.* **2020**, *17*, 199–223.
47. Kock, N.; Hadaya, P. Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Inf. Syst. J.* **2018**, *28*, 227–261.
48. Leguina, A. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*; Taylor & Francis: Abingdon, UK, 2015.
49. Kivivuori, J. *Discovery of Hidden Crime: Self-Report Delinquency Surveys in Criminal Policy Context*; Oxford University Press: Oxford, UK, 2011.
50. Baianstovu, R.Å.; Lundqvist, M.; Håkansson, S.Y. *Honour, Violence and Gender—An International Research Review*; Swedish Secretariat for Gender Research: Gothenburg, Sweden, 2025.
51. Chin, W.W. The partial least squares approach to structural equation modeling. In *Modern Methods for Business Research*; Psychology Press: New York, NY, USA, 1998; pp. 295–336.
52. Hair, J.F.; Risher, J.J.; Sarstedt, M.; Ringle, C.M. When to use and how to report the results of PLS-SEM. *Eur. Bus. Rev.* **2019**, *31*, 2–24.
53. Kiriş, A. Mediated Gendered Invisibility: Exploring Educated Young Women's Experiences of Online Sexual Harassment and Coping Strategies in Turkey. Master's Thesis, Middle East Technical University, Ankara, Turkey, 2024.
54. Kumar, A. Customer misbehavior in AI-enabled services the Motivations and Perceived Associationiveness of Online Counter-speech Writing and the Potential for AI-Mediated Assistance. Ph.D. Thesis, Virginia Tech, Blacksburg, VA, USA, 2024.
55. Hanafiah, M.H. Formative vs. reflective measurement model: Guidelines for structural equation modeling research. *Int. J. Anal. Appl.* **2020**, *18*, 876–889.
56. van Amelsvoort, L.G.; Fleuren, B.P.; Kant, I.J. Improving measurement models in clinical epidemiology: Time to move beyond the inherent assumption of an underlying reflective measurement model. *J. Clin. Epidemiol.* **2020**, *118*, 119–123.
57. Costa, D.S.J. Reflective, statistical, and composite indicators of quality of life: A conceptual or an empirical distinction? *Qual. Life Res.* **2015**, *24*, 2057–2065.
58. Loughran, J.J. *Developing Reflective Practice: Learning About Teaching and Learning Through Modelling*; Routledge: London, UK, 2002.
59. Bézivin, J.; et al. Reflective model driven engineering. in International Conference on the Unified Modeling Language. In *UML 2003—The Unified Modeling Language, Modeling Languages and Applications: Proceedings of the 6th International Conference, San Francisco, CA, USA, 20–24 October 2003*; Springer: Berlin/Heidelberg, Germany, 2003.
60. Henseler, J.; Ringle, C.M.; Sarstedt, M. A new criterion for assessing discriminant validity in variance-based structural equation modeling. *J. Acad. Mark. Sci.* **2015**, *43*, 115–135.
61. Fornell, C.; Larcker, D.F. Evaluating structural equation models with unobservable variables and measurement error. *J. Mark. Res.* **1981**, *18*, 39–50.
62. Becker, J.-M.; Cheah, J.-H.; Gholamzade, R.; Ringle, C.M.; Sarstedt, M. PLS-SEM's most wanted guidance. *Int. J. Contemp. Hosp. Manag.* **2022**, *35*, 321–346.
63. World Medical Association. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research on Human Beings. *JAMA* **2013**, *310*, 2191–2194.
64. Sarstedt, M.; Ringle, C.M.; Hair, J.F. Partial least squares structural equation modeling. In *Handbook of Market Research*; Springer: Cham, Switzerland, 2021 p. 587–632.

65. Henry, N.; Powell, A. Embodied harms: Gender, shame, and technology-facilitated sexual violence. *Violence Against Women* **2015**, *21*, 758–779.
66. Eckstein, J.J. Understanding darkness: Age, sex, and tech-proficiency in knowledge and perceptions of technology-mediated abuse. *J. Comput.-Mediat. Commun.* **2022**, *27*, zmac013.
67. Zhao, X.; Tan, H.; Zhang, Y.; Mattila, A.S.; Hua, N. Customer misbehavior in AI-enabled services. *Int. J. Contemp. Hosp. Manag.* **2025**, *37*, 3965–3982.
68. Majeed, A.; Rasheed, A. Green banking adoption in Pakistan: The UTAUT model augmented with green consumption values, environmental concerns, and collectivism culture. *J. Environ. Plan. Manag.* **2025**, *68*, 2883–2916.
69. Omweri, F.S. A systematic literature review of e-government implementation in developing countries: Examining urban-rural disparities, institutional capacity, and socio-cultural factors in the context of local governance and progress towards SDG 16.6. *Int. J. Res. Innov. Soc. Sci.* **2024**, *8*, 1173–1199.
70. Porto, M.; Golubeva, I.; Byram, M.J.S.S. Ethical Challenges in Intercultural Citizenship Education with ‘Difficult Topics’ in the World Language Classroom and Beyond. *Soc. Sci.* **2025**, *14*, 135.
71. Golubeva, I. *Diversity, Equity, Inclusion and Intercultural Citizenship In Higher Education*; Cambridge University Press: Cambridge, UK, 2025.
72. Barrett, M.; Golubeva, I. From intercultural communicative competence to intercultural citizenship: Preparing young people for citizenship in a culturally diverse democratic world, in Intercultural learning in language education and beyond. In *Intercultural Learning in Language Education and Beyond: Evolving Concepts, Perspectives and Practices*; Multilingual Matters: Bristol, UK, 2022; pp. 60–83.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.