Int J Ethics Soc 2026; 7(4): 65-78

doi 10.22034/ijethics.7.4.363

INTERNATIONAL JOURNAL OF
Ethics & Society

Original Article

Moral Disengagement and Cyberbullying among Iranian EFL Learners: Moral Identity as Mediator and Artificial Intelligence Self-Efficacy as Moderator

Seyyed Mohammad Reza Adel [1], Zahra Bayat[2*], Farhad Ghorbandordinejad[3]

1. Department of English, Faculty of Letters and Humanities, Ferdowsi University of Mashhad, Mashhad, Iran.
2. Department of English Language and Literature, Faculty of Literature and Humanities, Hakim Sabzevari University, Sabzevar, Iran.
3. Department of ELT, Faculty of Education, Baskent University, Ankara, Turkey.

Corresponding Author: Zahra Bayat, Department of English Language and Literature, Faculty of Literature and Humanities, Hakim Sabzevari University, Sabzevar, Iran. E-mail: bayat.zahra.1988@gmail.com

Received 31 Jun 2025 **Accepted** 28 Jul 2025 **Online Published** 04 Jan 2026

Abstract

Introduction: While considerable research has explored cyberbullying (CB) and its related factors, including moral disengagement (MD), moral identity (MI), and individual characteristics, the interplay of these elements remains underexplored, particularly among Iranian English as a Foreign Language (EFL) learners.
Material and Methods: Through proposing a new moderated-mediation model, this study investigated the mediating role of MI in the relationship between MD and CB, and the moderating role of artificial intelligence self-efficacy (AISE) in these connections. Applying a stratified random sampling, data were collected from 600 Iranian EFL learners using four standard questionnaires.
Results: The findings suggest that MI mediates the connection between MD and CB, while AISE moderates these relationships, highlighting the importance of fostering MI and responsible AI tool use in EFL contexts to mitigate CB.
Conclusion: The present investigation enhanced students' awareness of CB and fostered a culture of responsible digital citizenship, ultimately contributing to safer online learning environments. The implications suggested interventions and educational strategies designed to cultivate a more ethical and digitally responsible learning environment, finally reducing the prevalence of CB among EFL learners.
Keywords: *moral disengagement (MD), cyberbullying (CB), moral identity (MI), artificial intelligence self-efficacy (AISE)*

How to Cite: Adel S.M.R., Bayat Z, Ghorbandordinejad F. Moral disengagement and cyberbullying among iranian efl learners: moral identity as mediator and artificial intelligence self-efficacy as moderator, Int J Ethics Soc. 2026; 7(4): 65-78. doi: 10.22034/ijethics.7.4.363

INTRODUCTION

The digital era has facilitated communication and learning opportunities but has also increased the risks of negative online behaviors, particularly CB, which describes the use of digital platforms to harass or intimidate others and has become a global issue among adolescents [1-2]. Unlike traditional bullying, which occurs in physical spaces, CB follows victims into their private lives through online platforms and mobile devices, with anonymity and ease of access intensifying the problem. The psychological consequences of CB are severe, often resulting in distress, depression, and even suicidal ideation. This highlights the urgent need for comprehensive interventions and awareness programs, especially in schools [3].

CB is particularly relevant in EFL settings, where students might engage in it for fun or to express emotions, sometimes as language play or exaggeration. Students with lower English proficiency are particularly vulnerable to both perpetrating and experiencing CB, necessitating targeted interventions for this group. The rise of social media and technology has made CB more common, making it essential to understand its underlying causes [4, 5].

A key psychological concept for understanding CB is MD, which involves ignoring moral standards to justify harmful actions [6-9]. Research consistently demonstrates a positive association between MD and CB perpetration, with MD reducing feelings of guilt and self-accusation among perpetrators [7, 8]. Factors such as low agreeability, distracted parenting, and problematic social media use are linked to higher MD and increased CB [10-12]. Additionally, dark triad personality characteristics (traits like narcissism and psychopathy) were directly related to CB, with MD mediating this relationship.

On the other hand, MI, which refers to how much individuals define themselves by their moral values, serves as a protective factor against CB. A robust MI can moderate the impact of MD on CB, reducing the likelihood of such behaviors. Therefore, interventions that promote MI are crucial for combating CB. MI also moderated the effects of anger on CB and could be influenced by factors such as friends' MI and school climate. Childhood psychological maltreatment is linked to increased CB and MD, highlighting the broader implications of these constructs [13,14].

Another important factor is AISE, which is the belief in one's ability to effectively use AI technologies. AISE was positively correlated with trust in automated technology and positive attitudes toward AI. AISE likely moderates the relationship between MD and CB, as those with high AISE may be more aware of the consequences of their online actions and more likely to use AI responsibly [15-17]. In regulated digital environments like Iran, higher AISE was associated with greater adoption of digital tools such as VPNs, which might mediate CB dynamics. Using AI to combat CB raises concerns about youth perceptions and rights, emphasizing the need for their involvement in designing these interventions. Self-efficacy beliefs are crucial in mediating problem-solving behaviors, further underscoring the importance of AISE.

Despite extensive research on CB and its contributing factors, including MD and MI, there is still a lack of understanding of how these factors interact among Iranian EFL learners. While studies have explored MD's role in CB across cultures [18-20] and the protective effects of MI, few have studied the combined effect of MD and MI in the Iranian EFL setting. The potential moderating role of AISE, despite its correlation with technology acceptance [16, 17], is also underexplored in relation to CB, particularly within this specific cultural and educational setting.

Cultural factors also play a role in CB dynamics. In Iran, issues like language anxiety, cultural norms, and digital skills influence online behavior. Iran's strict internet policies and collectivist culture can foster environments where cyber aggression becomes a coping mechanism and bystander inaction is more common compared to Western contexts [21].

Social Cognitive Theory offers the theoretical framework for this investigation, suggesting that individuals can engage in harmful behaviors like CB by changing their moral standards through MD. MD allows for the justification of harmful actions, especially in anonymous online environments. Including MI as a mediator emphasizes how self-concept affects moral engagement, while AISE as a moderator suggests that confidence in using AI responsibly can reduce the negative effects of MD on CB [22-24]. This framework aligns with the theory's emphasis on how personal factors, behavior, and environmental elements collectively shape moral decision-making and conduct.

Given the relationships found in previous research [1-30], the present study investigates a model where MI mediates the link between MD and CB among Iranian EFL learners, and AISE moderates the relationships among MD, MI, and CB. The model posits that MI mitigates the effects of MD on CB, while high AISE may weaken the MD-CB link and strengthen MI's protective influence. This approach addresses gaps in the literature by examining the complex interplay of MD, MI, and AISE in predicting CB within a specific cultural and educational context, providing a novel understanding of the factors underlying CB among Iranian EFL learners.

Hypothesis 1: Moral identity (MI) will mediate the relationship between moral disengagement (MD) and cyberbullying (CB) among Iranian EFL learners.

Hypothesis 2: AI self-efficacy (AISE) will moderate the relationship between moral disengagement (MD) and moral identity (MI) among Iranian EFL learners, such that the relationship is weaker for those with high AISE.

Hypothesis 3: AI self-efficacy (AISE) will moderate the relationship between moral identity (MI) and cyberbullying (CB) among Iranian EFL learners, such that the relationship is weaker for those with high AISE.

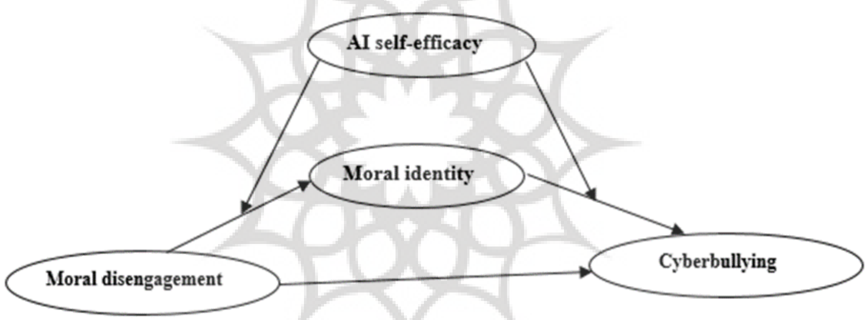

Figure 1: The Proposed Moderated-Mediation Model

MATERIAL AND METHODS

This research utilized a quantitative, cross-sectional design to investigate the relationships between MD, MI, AISE, and CB among Iranian EFL learners. Data were collected through a survey questionnaire administered to a sample of 600 EFL learners across five provinces in Iran. The questionnaire comprised established scales measuring MD, CB experiences, MI, and AISE. Following data collection, statistical analyses, including mediation and moderation analyses, were conducted using appropriate software to test the hypothesized relationships. The subsequent sections will provide detailed information on the participants involved in the study, the specific instruments used, the data collection procedure, and the data analysis techniques employed [31, 32].

Participants

The current research project was conducted across five distinct provinces of Iran: Tehran, Razavi Khorasan, Ardabil, Golestan, and Khuzestan. Participants were selected through a stratified random sampling method, ensuring that the sample accurately represented the diverse stratified population of learners in Iran. Data collection involved a total of 600 EFL learners, encompassing both male and female participants. Table 1 presents the demographic information of the participants involved in this study. As illustrated in the table, the number of female participants (N=381) exceeded that of male learners (N=219).

Table 1: Demographic Information

Participants	Frequency	Percent
Female	381	63.5
Male	219	36.5
Total	600	100

Instruments

Moral Disengagement Scale

To evaluate MD among participants, researchers used the Online Moral Disengagement Scale, developed by [33]. The mentioned tool consists of 8 items, each capturing a distinct mechanism of MD, such as rationalizing online behaviors or downplaying their severity. Participants respond based on a five-point Likert scale, from "strongly disagree" to "strongly agree." The mentioned tool is reliable, with a Cronbach's alpha of 0.81, and is effective in assessing MD mechanisms, helping to understand how individuals justify harmful online behaviors. For example, "I often justify my online actions by comparing them to worse behaviors" illustrates rationalization, a key mechanism of MD. This scale supported the study of MD in online contexts, where mechanisms like anonymity and disinhibition could facilitate harmful behaviors [33].

Cyberbullying Questionnaire

To assess CB experiences among Iranian EFL learners, researchers used the Cyber-Bullying/Victimization Experience Questionnaire consisting of 24 items and developed by [34]. This tool evaluates the prevalence and impact of CB, employing a five-point Likert scale from "strongly disagree" to "strongly agree." This tool demonstrated strong internal consistency with a high Cronbach's alpha, indicating robust reliability. In their research, its validity was supported by a two-factor model confirmed through Confirmatory Factor Analysis (CFA), with good fit indices such as a Comparative Fit Index (CFI) of 0.96 and a Root Mean Square Error of Approximation (RMSEA) of 0.088. For example, "I have been mocked online by someone" illustrates a direct CB experience, while "Someone has spread rumors about me online" represents an indirect form of CB. This validated tool aids in understanding and mitigating the effects of CB among adolescents [34].

Moral Identity Questionnaire

To examine the MI of EFL learners, researchers used the Moral Identity Questionnaire (MIQ), having 20 items and developed by [24]. This instrument assesses two key components of MI: moral self, which reflects how important moral principles are to an individual's self-concept, and moral integrity, which focuses on acting in accordance with one's moral principles. For example, "Being honest is a fundamental part of who I am" illustrates moral self, while "I always try to act in ways that align with my moral values" demonstrates moral integrity. The MIQ is a reliable tool, using a 6-point Likert scale and showing strong internal consistency with a Cronbach's alpha coefficient of 0.87 [24].

Artificial Intelligence Self-Efficacy Scale

To analyze participants' AISE, this study applied the Artificial Intelligence Self-Efficacy Scale developed and validated by [35]. The AISES is a 22-item instrument designed to measure the extent to which an individual feels capable of using AI technologies and products proficiently. Participants respond to the items based on a seven-point Likert scale, classifying from "strongly disagree" to "strongly agree." Four various components make up the scale: assistance, anthropomorphic interaction, comfort with AI, and technological skills. For assistance, an example item might be, "I believe AI tools can effectively help me complete tasks more efficiently." Regarding anthropomorphic interaction, a sample item could be, "I feel comfortable communicating with AI systems that mimic human conversation." To assess comfort with AI, a representative item might be, "I am not worried about using AI in my daily life." Finally, concerning technological skills, an example item could be, "I have the essential technical skills to effectively use AI applications." The AISES demonstrated good fit, reliability, convergent validity, discriminant validity, content validity, and criterion-related validity [35]. Moreover, the scale exhibited the validity through a positive correlation with motivated learning behaviors, making it a useful tool for assessing AISE.

Procedure

To gather the necessary data for this study, a total of 600 EFL learners were invited to participate by completing the relevant scales after providing demographic information. Prior to participation, all learners gave informed consent to ensure adherence to ethical standards throughout the research process. Measures were taken to ensure participant confidentiality and anonymity. Ethical approval was obtained from both the Hakim Sabzevari University Ethics Committee and the Ministry of Education's Ethics Committee across five Iranian provinces: Tehran, Razavi Khorasan, Ardabil, Golestan, and Khuzestan, prior to the data collection process.

Our study acknowledged the importance of ethical approval and informed consent but recognized a need for enhanced participant protection measures. Given that recalling experiences related to CB can evoke emotional distress and anxiety [29], we proposed integrating debriefing sessions and offering psychological support to participants who may experience distress. To mitigate risks further, we ensured strict anonymity in survey responses and employed indirect questioning techniques to safeguard participants' well-being [36].

Data collection was conducted over several weeks using both online and paper-based questionnaires to maximize accessibility for participants. Online surveys were distributed via Google Forms, with links shared through the Shad school messaging system. Paper surveys were available for those preferring traditional methods or lacking reliable internet access. Participants received clear instructions verbally and in writing, emphasizing the importance of honest responses and confidentiality. They were also informed of their right to withdraw from this project at any time, with assurances of maintaining procedural autonomy and avoiding negative consequences. Ethical authorization was secured from the Ministry of Education in Iran, and permission was secured from the identified EFL learners as well as their guardians. Anonymity was maintained by assigning unique identifiers to each questionnaire, allowing responses to be analyzed while preserving privacy.

To enhance the feasibility of collecting data from 600 Iranian EFL learners, our study implemented strategies to minimize dropout rates. Recognizing that lengthy and complex surveys can lead to participant fatigue [37], we employed engagement tactics such as offering small incentives like digital certificates or educational resources, and utilizing shorter, more interactive survey formats [38]. Additionally, we scheduled follow-ups with non-respondents to mitigate attrition and improve data quality. By optimizing survey design and incorporating participant-friendly features, we aimed to maintain high participation rates throughout the study.

The study utilized several tools to assess different aspects of the participants. The Moral Disengagement Scale, developed by [33], was used to evaluate MD. The mentioned instrument by [34] assessed CB experiences. The MIQ by [24] evaluated MI. Lastly, the Artificial Intelligence Self-Efficacy Scale by [35] measured participants' perceived self-efficacy in using AI technologies.

Data Analysis

To explore the trends and characteristics of each variable, a detailed descriptive statistical analysis was conducted. This involved calculating key statistics such as the mean, standard deviation, and the range of item averages, including minimum and maximum values. Additionally, a correlation analysis was applied to investigate the relationships between MD, CB, MI, and AISE. To provide deeper insight into these relationships, AMOS 26 and the PROCESS macro in SPSS version 26 were used for mediation and moderation analyses. This analysis aimed to uncover the complex interactions between the variables, providing insights into their interdependencies [31,39].

Prior to the main analysis, CFA was conducted to evaluate the convergent and discriminant validity of the measurement instruments used in the study. Furthermore, the internal consistency of these measures was evaluated using Cronbach's alpha, with a reliability threshold of 0.7, as suggested by [32]. This rigorous approach ensured that the measurement tools were both valid and reliable, establishing a solid foundation for subsequent analyses. In this regard, [39] used the MLE approach to ensure methodological precision. The model fit indices met or exceeded standard thresholds (e.g., the Goodness-of-Fit Index or GFI, the Normed Fit Index or NFI, the Incremental Fit Index or IFI, CFI > 0.9; RMSEA < 0.06), supporting the study's objectives with a comprehensive analytical framework.

By employing these statistical methods, the study sought to guarantee the precision and dependability of the results, thereby contributing to a deeper knowledge of the variables under investigation. The combination of descriptive statistics, correlation analysis, mediation, and moderation analyses provided a comprehensive framework for exploring the complex dynamics between MD, CB, MI, and AISE among Iranian EFL learners.

RESULTS

At the outset of the analysis, Harman's single-factor test was employed to assess the potential for common method bias (CMB) in self-report data, following the approach suggested by [36]. The results showed that the primary factor explained just 35.41% of the total variance, suggesting that CMB is unlikely to be a significant concern. This outcome indicated that each measure effectively captured unique aspects of its respective construct, without being unduly influenced by biases related to social desirability.

The present examination validated its scales by conducting the CFA analysis on each set of items. The reports indicated that the MD Scale, with 8 items, achieved a CFI of 0.946, a GFI of 0.957, and an RMSEA of 0.048. The CB Questionnaire, comprising 24 items, showed a CFI of 0.971, a GFI of 0.989, and an RMSEA of 0.045. Meanwhile, the MI Questionnaire, consisting of 20 items, had a CFI of 0.968, a GFI of 0.947, and an RMSEA of 0.049. Additionally, the Artificial Intelligence Self-Efficacy Scale, with 22 items, demonstrated a CFI of 0.973, a GFI of 0.952, and an RMSEA of 0.041. These results generally align with the criteria for good model fit (CFI > 0.9, GFI > 0.9, RMSEA < 0.06), supporting the validity of all scales in structural equation modeling (SEM) contexts (Table 2).

Table 2: Validation of indices for each scale

Scales	number of items	CFI	GFI	RMSEA

Moral Disengagement Scale	8	0.946	0.957	0.048
Cyberbullying Questionnaire	24	0.971	0.989	0.045
Moral Identity Questionnaire	20	0.968	0.947	0.049
Artificial Intelligence Self-Efficacy Scale	22	0.973	0.952	0.041
Total	74			
Acceptable fit indices		>0.9	>0.9	<0.06

Preliminary Analysis

The table data can be statistically reported as follows: The dataset included four components, MD, CB, MI, and AISE, with each having a sample size of 600. The skewness values ranged from -0.1 to 0.2, indicating relatively symmetrical distributions. Kurtosis values were between 1.5 and 1.9, suggesting slightly platykurtic distributions. The mean scores varied significantly across components, with MD at 18.4, CB at 65.7, MI at 78.3, and AISE at 89.2. Standard deviations are 2.2 for MD, 5.7 for CB, 4.6 for MI, and 6.5 for AISE. Correlations between components show significant relationships, with notable correlations between MD and CB (r = -0.7), MD and MI (r = 0.6), and Moral Disengagement and AISE (r = -0.4). These findings align with the principles outlined by [32] for interpreting statistical data in psychological research.

Table 3: Descriptive statistics

Components	N	Skewness	Kurtosis	*Mean*	*SD*	1	2	3
Moral Disengagement (MD)	600	0.2	1.5	18.4	2.2	-		
Cyberbullying (CB)	600	0.1	1.6	65.7	5.7	-0.4**	-	
Moral Identity (MI)	600	-0.1	1.8	78.3	4.6	0.6**	-0.3**	-
AI Self-Efficacy (AISE)	600	0	1.9	89.2	6.5	-0.7**	-0.5**	0.4**

SEM Analysis

Mediation Analysis (Hypothesis 1)

Addressing the first hypothesis, the SEM model's overall fit to the data was assessed using several indices, as presented in Table 4. The model demonstrated acceptable fit based on the established criteria: The P(chi-square) value was 0.17 (> 0.05), indicating a non-significant discrepancy between the actual and estimated covariance matrices. The GFI was 0.945 (> 0.9), the NFI was 0.973 (> 0.9), the Incremental Fit Index (IFI) was 0.979 (> 0.9), and the CFI was 0.936 (> 0.9). Additionally, the RMSEA was 0.042 (< 0.06). Regarding [39], these indices collectively suggested that the proposed model adequately represented the interconnections among the variables.

Table 4: Goodness of fit indices

	P(chi-square)	GFI	NFI	IFI	CFI	RMSEA
Acceptable fit	>0.05	>0.9	>0.9	>0.9	>0.9	<0.06
Model	0.17	0.945	0.973	0.979	0.936	0.042

Furthermore, Table 5 presents the direct and indirect effects from the SEM analysis. Regarding direct effects, MD had a significant negative effect on MI (B = -0.21, p < 0.001, 95% CI [-0.31, -0.41]). MI also exhibited a considerable negative effect on CB (B = -0.42, p < 0.001, 95% CI [-0.47, -0.79]). Also, MD had a meaningful direct effect on CB (B = 0.20, p < 0.001, 95% CI [0.30, 0.60]). The indirect effect, representing the mediation of MI, was also significant (B = 0.15, p < 0.001, 95%

CI [0.08, 0.15]), indicating that MD influences CB through its effect on MI. In summary, all the effects examined were statistically significant and supported, as registered in the "Decision" column of the table.

Table 5: Direct and indirect effects of SEM analysis

Path	*B*	SE	95% CI [Lower bound; Upper bound]	*P*	Decision
Direct Effects					
MD→MI	-0.21	0.05	[-0.31; -0.41]	<0.001	Supported
MI→CB	-0.42	0.05	[-0.47; -0.79]	<0.001	Supported
MD→CB	0.20	0.05	[0.30; 0.60]	<0.001	Supported
Indirect Effects					
MD→MI→CB	0.15	0.02	[0.08; 0.15]	<0.001	Supported
Total Effects					
MD→CB	0.63	0.05	[0.81; 0.91]	<0.001	Supported

Moderation Analysis (Hypotheses 2 & 3)

Regarding [31] and based on Figure 2, which illustrates the moderation of AISE in the relationship between MD and MI (Hypothesis 2), the data suggested that AISE moderated this association. As shown in Figure 2, the interaction term between MD and AISE significantly moderated the effect of MD on MI ($b = -0.08$, $p < 0.01$, 95% CI [-0.12, -0.04]), suggesting that AISE moderated the direct link between MD and MI. Specifically, for Iranian EFL learners with high AISE, the relationship between MD and MI appeared to be weaker compared to those with low AISE. This was evidenced by the less steep slope of the "High AISE" line relative to the "Low AISE" line, indicating that as MD increases, the corresponding increase in MI is less pronounced for individuals with high AISE.

Concerning Figure 3 and the recommendations by [31], Hypothesis 3, which posited that AISE moderates the link between MI and CB among Iranian EFL learners, appeared to be supported. As shown in Figure 3, the interaction term between MI and AISE significantly moderated the effect of MI on CB ($b = -0.15$, $p < 0.001$, 95% CI [-0.08, -0.22]), suggesting that AISE moderated the direct link between MI and CB. The dashed line, representing high AISE, showed a less steep negative slope than the solid line, representing low AISE. This indicated that as MI increases, CB decreases more rapidly for participants with high AISE compared to those with low AISE. To explain explicitly, high MI lessens the impact of MD on CB. Also, high AISE weakens the MD→MI link and enhances the MI→CB protective effect**.** In other words, the relationship between MI and CB is weaker when AISE is high. This supports the idea that high AISE strengthens MI's ability to protect against CB. Thus, the relationship between MI and CB was weaker for those with high AISE, suggesting that high AISE enhances the protective effect of MI against CB, supporting the hypothesis as originally predicted.

Figure 2: The moderation of AI self-efficacy in the relationship between moral disengagement and moral identity (H2)

Figure 3: The moderation of AI self-efficacy in the relationship between moral identity and cyberbullying (H3)

DISCUSSION

This study investigated the intricate relationships between MD, MI, AISE, and CB among Iranian EFL learners. The findings strongly supported the proposed moderated-mediation model, emphasizing the crucial role of MI in mediating the connection between MD and CB, with AISE acting as a significant moderator in these dynamics.

The first hypothesis predicted that MI would mediate the relationship between MD and CB, and this was confirmed. This corresponds with earlier findings underscoring the protective role of a strong MI in mitigating CB behaviors [9,14]. Iranian EFL learners who defined themselves through moral traits and values are less likely to engage in CB, even when prone to MD. This supports studies showing that MI helps reduce the negative effects of MD on CB [15]. In line with [27], this study highlighted the broader impact of MI on online behaviors, as MI partially mediates the link between moral perfectionism and online prosocial behavior. Furthermore, the finding that self-other overlap helps predict positive online behavior through empathy and MI shows how important MI is for good online interactions [28].

The second hypothesis stated that AISE would moderate the relationship between MD and MI, and this was also confirmed. This indicated that the negative impact of MD on MI was lessened when individuals had high confidence in their AI abilities. EFL learners with high AISE might be more aware of the ethical implications of their online actions, strengthening their MI even when facing situations that might otherwise lead to MD. This supports findings that show AISE is linked to a greater tendency to trust automated technology and have a positive attitude toward AI [16].

The third hypothesis proposed that AISE would moderate the bond between MI and CB, and this was confirmed. As shown in Figure 3, the interaction term between MI and AISE significantly moderated the effect of MI on CB, indicating that AISE influences the strength of this relationship. This means that as MI increases, CB decreases more modestly among students with high AISE compared to those with low AISE. When AISE is high, the relationship between MI and CB becomes weaker, supporting the idea that AISE can dampen the influence of MI on CB in some pathways. At the same time, high AISE supports the protective role of MI through other mechanisms, such as reducing MD's impact on MI, which helps explain the complex moderating role of AISE.

Discussing these findings using Social Cognitive Theory helps us understand how personal, behavioral, and environmental factors interact in CB. According to the stated framework, EFL learners learn and adopt behaviors through observing others and their own experiences [22]. The moderating role of AISE suggested that confidence in using AI responsibly could enhance MI and reduce the likelihood of CB by promoting ethical online behaviors. This aligns with the theory's focus on how personal factors (like MD and MI) and environmental influences (like social norms and technology use) shape behavior [22,40]. By incorporating AI education into moral development programs, educators can foster a culture of responsible technology use, which is crucial for mitigating CB among EFL learners. This approach supports the theory's principles by addressing both personal and environmental factors to prevent bullying behaviors [6,41].

Compared to previous research, this study made a significant contribution by demonstrating the moderation of AISE in the relationship between MI and CB, particularly within the Iranian EFL context. While prior research has explored the role of MD in CB across different cultures [18-20] and the protective effects of MI [9,14], this study added novel insights into the moderating influence of AISE. The research also addressed the gap identified in the literature review by discovering the combined effect of MI and AISE within a novel moderated-mediation model, building upon the work of [5] in identifying CB actions and proposing anti-CB frameworks.

Moreover, the findings by [42] on the "responsibility gap" in AI relate to our outputs by emphasizing the need for human agency in ethical technology use. Another study identified various moral damages caused by students' lifestyles influenced by cyberspace, including social, educational-cultural, political, security, psychological, and physiological harms. Addressing these identified damages is essential for mitigating their impact and promoting responsible use of cyberspace [43]. Also, the study reveals that Allameh Tabatabai's principles of moral education emphasize knowledge-based education, guardianship, and the pursuit of monotheistic values, self-knowledge, and human harmony. The application of these principles in cyberspace aims to awaken inner restraint, highlighting the crucial role of guides in this process [44].

It should be noted that EFL learners encountered unique challenges that could lead to CB behaviors, influenced by language barriers and online communication norms [5]. Students with lower English proficiency might engage in CB as a defense mechanism to hide insecurities [14]. Peer pressure to use slang or aggressive language in online forums could also foster hostile interactions [9].

Concerning [5], the findings suggested that effective interventions should focus on enhancing both MI and AISE among EFL learners. By strengthening moral values and increasing confidence in responsible AI technology use, EFL learners could be empowered to resist CB and foster a safer online environment. Addressing issues of trust and fear related to AI, as highlighted by [17] and [1], was crucial for promoting responsible AI use and preventing CB among Iranian EFL learners. This aligns with the emphasis on involving youth in designing AI interventions to meet their needs, rights, and ethical considerations [29].

CONCLUSION

The current project provided compelling evidence for the intricate relationships between MD, MI, AISE, and CB among Iranian EFL learners. The outputs confirmed the mediation of MI in the connection between MD and CB. Significantly, the results demonstrated the moderating influence of AISE on both the link between MD and MI, and the association between MI and CB, highlighting the protective role of AISE. This novel moderated-mediation model offered a nuanced understanding of the factors influencing CB in this specific context, going beyond the existing literature by combining these elements in a single, comprehensive framework. The support for all three hypotheses underscores the importance of fostering both MI and AISE to mitigate CB among Iranian EFL learners.

While this study provided valuable insights, it's valuable to acknowledge its limitations, particularly those specific to the research design and sample. The cross-sectional nature of the study hindered drawing causal inferences. We could only observe associations between the variables at a single point in time, not determine how changes in one variable might lead to changes in others. Additionally, reliance on self-report questionnaires might be subject to social desirability bias. Participants might underreport CB perpetration or MD due to concerns about how their answers might be perceived. The current sample was limited to Iranian EFL learners, which reduced the generalizability of the findings to diverse populations or cultural backgrounds. EFL learners in other countries or even students in other fields of study in Iran might exhibit different patterns. In addition, the study used a specific measure of AISE, which may need to be updated in future research to reflect current AI tools and applications. Lastly, the quantitative design did not capture the nuanced experiences and perspectives of EFL learners involved in CB.

Building upon the results and addressing the shortcomings of this research journey, several suggestions for further investigations are proposed. Conducting longitudinal studies would allow for a better understanding of the causal relationships between MD, MI, AISE, and CB over time. Employing a mixed-methods approach, integrating quantitative data with qualitative data (e.g., interviews, focus groups), would result in a better understanding of the complex dynamics of CB and the lived experiences of EFL learners involved in these behaviors. Replicating this study in different cultural contexts would assess the generalizability of the findings and identify potential cultural variations in the associations between the variables. Experimental studies could test the effectiveness of interventions organized to promote MI and enhance AISE in reducing CB behaviors. Additionally, investigating the role of AISE in relation to specific AI applications used in online interactions would provide a more nuanced understanding of how AISE influences online behavior in specific contexts. Exploring other potential moderators (e.g., online disinhibition, empathy) and mediators (e.g., social support, ethical awareness) that might influence the relationships between MD, MI, AISE, and CB is also recommended. Finally, developing and evaluating AI ethics education programs could promote responsible AI use and prevent CB by addressing the ethical use of AI and teaching students how to use AI technologies responsibly.

To enhance the understanding of CB in the context of language learning, it is crucial to incorporate a cross-cultural perspective into the future research. While this study provides valuable insights from a local Iranian context, comparisons with other cultural backgrounds, such as those in East Asia or Western Europe, could significantly enrich the discourse. Exploring how different cultural norms, values, and educational practices influence the dynamics of MI, AISE, and CB would offer a more comprehensive understanding of these relationships. Such comparative analyses could reveal unique patterns and challenges faced by language learners in various settings, ultimately contributing to the development of more effective, culturally sensitive interventions aimed at reducing CB and promoting positive online behavior across diverse educational environments.

Future studies could explore the implementation of targeted interventions, such as integrating AI-powered behavioral monitoring systems within EFL classrooms or developing gamified modules focused on digital responsibility [29]. These innovative approaches could enhance students' awareness of CB and foster a culture of responsible digital citizenship, ultimately contributing to safer online learning environments. This journey contributes to the developing field of knowledge on CB by highlighting the importance of MI and AISE in mitigating CB among Iranian EFL learners. The explorations underscore the need for interventions that foster both moral development and responsible AI technology use. By empowering EFL learners with strong moral values and the skills to navigate the online world safely and ethically, we can assist with creating a more positive and inclusive digital environment for all.

The findings of this study suggest that enhancing MI and AISE among Iranian EFL learners is crucial for mitigating CB behaviors. By fostering these competencies, educators can create a more supportive and ethical online learning environment. This can be achieved through targeted interventions, such as workshops and training sessions that focus on ethical decision-making and responsible technology use. Additionally, integrating discussions about the implications of AI in educational settings can empower students to navigate digital interactions more thoughtfully, ultimately reducing the incidence of CB.

Furthermore, the importance of a cross-cultural perspective in future research cannot be overstated. As the dynamics of MI, AISE, and CB may vary significantly across different cultural contexts, comparative studies could provide deeper insights into how these factors interact globally. This understanding can inform the development of culturally sensitive educational programs that address the unique challenges faced by EFL learners in various regions. By tailoring interventions to reflect the cultural values and norms of diverse populations, educators can enhance the effectiveness of strategies aimed at promoting positive online behavior and reducing CB, thereby contributing to a safer and more inclusive digital landscape for all learners.

ETHICAL CONSIDERATIONS

The researchers have addressed ethical concerns, including plagiarism, deliberate deception, data falsification or fabrication, duplicate publication or submission to multiple venues, redundancy, and other related issues.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests.

FUNDING DECLARATION

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

REFERENCES

1. Ghosh R, Malhotra M, Kumar N. Cyber bullying in the digital age: challenges, impact, and strategies for prevention. In Combating Cyberbullying with Generative AI (pp. 151-180). IGI Global Scientific Publishing. 2025. https://doi.org/10.4018/979-8-3373-0543-1.ch006
2. Sun Y X, Cao C H, Tang Z J, Huang F M, Zhong X B, Chen I H. Moral disengagement as mediator and guilt as moderator between cyber moral literacy and cyberbullying among late adolescents. Scientific Reports, 2025; 15(1): 43. https://doi.org/10.1038/s41598-024-83911-5
3. Shahzad K, Khan S A, Javeed A M D, Iqbal A. Factors influencing cyberbullying among citizens: a systematic review of articles published in refereed journals from 2010 to 2023. Global Knowledge, Memory and Communication. 2024. https://doi.org/10.1108/gkmc-11-2023-0422
4. Lin C S, Kuo Y F, Wang T Y. Trust and acceptance of AI caregiving robots: The role of ethics and self-efficacy. Computers in Human Behavior: Artificial Humans, 2025; 3: 100115. https://doi.org/10.1016/j.chbah.2024.100115
5. Hamuddin B, Rahman F, Pammu A, Sanusi Baso Y, Derin T. Cyberbullying among EFL students' blogging activities: Motives and proposed solutions. Teaching English with Technology, 2020; 20(2): 3-20. https://doi.org/10.22437/irje.v4i2.10202
6. Bussey K, Fitzpatrick S, Raman A. The role of moral disengagement and self-efficacy in cyberbullying. Journal of School Violence, 2015; 14(1): 30-46. https://doi.org/10.1080/15388220.2014.954045
7. Li W. Moral disengagement, herd mentality, moral identity, and empathy in cyberbullying roles. Journal of Aggression, Maltreatment & Trauma, 2023; 32(9): 1198-1220. https://doi.org/10.1080/10926771.2023.2222079
8. Lo Cricchio M G, García-Poole C, Te Brinke L W, Bianchi D, Menesini E. Moral disengagement and cyberbullying involvement: A systematic review. European Journal of Developmental Psychology, 2021; 18(2): 271-311. https://doi.org/10.1080/17405629.2020.1782186
9. Wang X, Yang L, Yang J, Wang P, Lei L. Trait anger and cyberbullying among young adults: A moderated mediation model of moral disengagement and moral identity. Computers in Human Behavior, 2017; 73: 519-526. https://doi.org/10.1016/j.chb.2017.03.073
10. Colella G M, Palermiti A L, Bartolo M G, Servidio R C, Costabile A. Problematic social media use, retaliation, and moral disengagement in cyberbullying and cybervictimization among Italian preadolescents: A moderated mediation model. International Journal of Bullying Prevention, 2024; 1-16. https://doi.org/10.1007/s42380-024-00265-2
11. Gao L, Li X, Wang X. Agreeableness and adolescents' cyberbullying perpetration: A longitudinal moderated mediation model of moral disengagement and empathy. Journal of Personality, 2023; 91(6): 1461-1477. https://doi.org/10.1111/jopy.12823
12. Wang X, Wang W, Qiao Y, Gao L, Yang J, Wang P. (2022). Parental phubbing and adolescents' cyberbullying perpetration: A moderated mediation model of moral disengagement and online disinhibition. Journal of Interpersonal Violence, 2022; 37(7-8): NP5344-NP5366. https://doi.org/10.1177/0886260520961877
13. Gholami M, Thornberg R, Kabiri S, Yousefvand S. From dark triad personality traits to digital harm: mediating cyberbullying through online moral disengagement. Deviant Behavior, 2025; 1-19. https://doi.org/10.1080/01639625.2025.2453445
14. Li H, Guo Q, Hu P. Moral disengagement, self-control and callous-unemotional traits as predictors of cyberbullying: a moderated mediation model. BMC Psychology, 2023; 11(1): 247. https://doi.org/10.1186/s40359-023-01287-z
15. Lyu W, Zhang J. The influence of childhood psychological maltreatment on mainland China college students' cyberbullying: The mediating effect of moral disengagement and the moderating effect of moral identity. Eurasia Journal of Mathematics, Science and Technology Education, 2017; 13(11): 7581-7590. https://doi.org/10.12973/ejmste/80302
16. Obenza B N, Baguio J S I, Bardago K M, Granado L, Loreco K C, Matugas L, ... & Caangay R B. The mediating effect of AI trust on AI self-efficacy and attitude toward AI of College Students. International Journal of Metaverse (IJM), 2024; 2(1). https://doi.org/10.54536/ijm.v2i1.2286
17. Montag C, Kraus J, Baumann M, Rozgonjuk D. The propensity to trust in (automated) technology mediates the links between technology self-efficacy and fear and acceptance of artificial intelligence. Computers in Human Behavior Reports, 2023; 11: 100315. https://doi.org/10.1016/j.chbr.2023.100315
18. Shohoudi Mojdehi A, Leduc K, Shohoudi Mojdehi A, Talwar V. Examining cross-cultural differences in youth's moral perceptions of cyberbullying. Cyberpsychology, Behavior, and Social Networking, 2019; 22(4): 243-248. https://doi.org/10.1089/cyber.2018.0339
19. Wang X, Lei L, Liu D, Hu H. Moderating effects of moral reasoning and gender on the relation between moral disengagement and cyberbullying in adolescents. Personality and individual differences, 2016; 98: 244-249. https://doi.org/10.1016/j.paid.2016.04.056
20. Zhao L, Yu J. A meta-analytic review of moral disengagement and cyberbullying. Frontiers in Psychology, 2021; 12: 681299. https://doi.org/10.3389/fpsyg.2021.681299
21. Dewaele J M, MacIntyre P D. Foreign language enjoyment and foreign language classroom anxiety: The right and left feet of the language learner. Positive Psychology in SLA, 2016; 215(236). https://doi.org/10.21832/9781783095360-010

22. Bandura A. Social cognitive theory: An agentic perspective. Annual Review of Psychology, 2001; 52(1): 1-26. https://doi.org/10.1146/annurev.psych.52.1.1
23. Bandura A. Social-learning theory of identificatory processes. Handbook of socialization theory and research, 1969; 213: 262. https://doi.org/10.1111/j.1939-0025.1963.tb01007.x
24. Black J E, Reynolds W M. Development, reliability, and validity of the Moral Identity Questionnaire. Personality and Individual Differences, 2016; 97: 120-129. https://doi.org/10.1016/j.paid.2016.03.041
25. Thornberg R, Daremark E, Gottfridsson J, Gini G. Situationally selective activation of moral disengagement mechanisms in school bullying: A repeated within-subjects experimental study. Frontiers in Psychology, 2020; 11: 1101. https://doi.org/10.3389/fpsyg.2020.01101
26. Wang X, Zhao F, Yang J, Lei L. (2021). School climate and adolescents' cyberbullying perpetration: A moderated mediation model of moral disengagement and friends' moral identity. Journal of Interpersonal Violence, 2021; 36(17-18): NP9601-NP9622. https://doi.org/10.1177/0886260519860089
27. Zeng P, Zhao X, Xie X, Long J, Jiang Q, Wang Y, ... & Wang P. Moral perfectionism and online prosocial behavior: The mediating role of moral identity and the moderating role of online interpersonal trust. Personality and Individual Differences, 2020; 162: 110017. https://doi.org/10.1016/j.paid.2020.110017
28. Dong Z, Wang R, Zhao Z, Xie K, Li A, Luo H, Li L. How self-other overlap shapes online altruism in adolescents: the role of empathy and moral identity. Psychology Research and Behavior Management, 2025; 331-340. https://doi.org/10.2147/prbm.s492811
29. Milosevic T, Verma K, Carter M, Vigil S, Laffan D, Davis B, O'Higgins Norman J. (2023). Effectiveness of artificial intelligence–based cyberbullying interventions from youth perspective. Social Media+ Society, 2023; 9(1): 20563051221147325. https://doi.org/10.1177/20563051221147325
30. Ferreira P C, Simão A V, Paiva A, Ferreira A. Responsive bystander behaviour in cyberbullying: A path through self-efficacy. Behaviour & Information Technology, 2020; 39(5): 511-524. https://doi.org/10.1080/0144929x.2019.1602671
31. Hayes A F, Rockwood N J. Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. Behaviour Research and Therapy, 2017; 98: 39-57. https://doi.org/10.1016/j.brat.2016.11.001
32. Tabachnick B G, Fidell L S, Ullman J B. Using multivariate statistics. 6th ed. Pearson. 2013. https://doi.org/10.1037/e676252012-088
33. Paciello M, Tramontano C, Nocentini A, Fida R, Menesini, E. The role of traditional and online moral disengagement on cyberbullying: Do externalising problems make any difference? Computers in Human Behavior, 2020; 103: 190-198. https://doi.org/10.1016/j.chb.2019.09.024
34. Nooripour R, Hosseinian S, Ghanbari N, Wisniewski P, Sikström S. Validity and reliability of Persian version of cyber-bullying/victimization experience questionnaire (CBVEQ) among Iranian adolescents. International Journal of Bullying Prevention, 2024; https://doi.org/10.1007/s42380-024-00211-2.
35. Wang Y Y, Chuang Y W. Artificial intelligence self-efficacy: Scale development and validation. Education and Information Technologies, 2024; 29(4): 4785-4808. https://doi.org/10.1007/s10639-023-12015-w
36. Podsakoff P M, MacKenzie S B, Podsakoff N P. Sources of method bias in social science research and recommendations on how to control it. Annual Review of Psychology, 2012; 63(1): 539-569. https://doi.org/10.1146/annurev-psych-120710-100452
37. Dillman D A, Smyth J D, Christian L M. Internet, phone, mail, and mixed-mode surveys: The tailored design method. 4th ed. USA: Wiley. 2014. https://doi.org/10.1002/9781394260645
38. Galesic M, Bosnjak M. Effects of questionnaire length on participation and indicators of response quality in a web survey. Public Opinion Quarterly, 2009; 73(2): 349-360. https://doi.org/10.1093/poq/nfp031
39. Meyers L S, Gamst G S, Guarino A J. Applied Multivariate Research: Design and Interpretation. Sage. 2008. https://doi.org/10.4135/9781071802687
40. Salmivalli C, Peets K. Bullying and victimization. In W. M. Bukowski, B. Laursen, & K. H. Rubin (Eds.), Handbook of peer interactions, relationships, and groups. 2nd ed. The Guilford Press. 2018. https://doi.org/10.1002/9781444390933.ch27
41. Swearer S M, Wang C, Berry B, Myers Z R. Reducing bullying: Application of social cognitive theory. Theory into practice, 2014; 53(4): 271-277. https://doi.org/10.1080/00405841.2014.947221
42. Schur E, Brouns A, Lee P. Ethical analysis of the responsibility gap in artificial intelligence. International Journal of Ethics and Society. 2025; 6(4): 1-10 https://doi.org/10.22034/ijethics.6.4.1
43. Khalkhali M, Niaz Azari K, Enayati T. Identifying moral damages caused by the lifestyle of students affected by cyberspace. International Journal of Ethics and Society, 2023; 5(2): 49-57. https://doi.org/10.22034/ijethics.5.2.69
44. Eghbalian T, Samadi M, Beheshti S. The pattern of application of moral education in cyberspace based on the view of Allameh Tabatabai. International Journal of Ethics and Society, 2022; 4(2): 3. https://doi.org/10.52547/ijethics.4.2.11