

Religious Moderation and Cyber Social Morality: An Internalization Study among Students

Roikhatul Jannah*, Saeful Bahri, Ahmad Hufron

Institut Agama Islam Bakti Negara Tegal, Indonesia

*Correspondence: ✉ raichah.jannah@gmail.com

Abstract

The rapid expansion of digital interaction among university students has created new challenges for sustaining religious moderation values in online environments. This study examines the process of internalizing religious moderation values among students of the Institut Agama Islam Bakti Negara (IAIBN) Tegal and its implications for the manifestation of their social morality in cyberspace. Amid the government's extensive efforts to strengthen religious moderation—particularly within Islamic Higher Education Institutions (PTKI)—the digital sphere has become a dominant arena of social interaction that presents unique psychological challenges. Employing a qualitative case-study approach, this research collected data through in-depth interviews, digital participatory observation (digital ethnography), and document analysis. The data were analyzed thematically by integrating the theoretical frameworks of value internalization and John Suler's Online Disinhibition Effect. The findings reveal that the internalization of religious moderation values at IAIBN Tegal—implemented through initiatives such as the *Rumah Moderasi* and curriculum integration—has been largely successful at the cognitive level. However, the manifestation of these values in students' social morality within cyberspace remains complex and ambivalent. Online disinhibition factors—such as dissociative anonymity, invisibility, and dissociative imagination—significantly shape students' behavior, often creating discrepancies between the values they espouse and the actions they display online. These dynamics may lead to benign disinhibition (positive openness) that aligns with moderation values, or to toxic disinhibition (aggressiveness and intolerance) that contradicts them. This study concludes that value internalization is contextual rather than absolute; the situational forces of cyberspace can temporarily override internalized values. The findings underscore the urgency for developing a "Digital Moderation Literacy" curriculum within PTKI—one that goes beyond theological knowledge transfer and equips students with metacognitive skills to consciously recognize and manage the effects of online disinhibition.

Article History

Received: 05-05-2024

Revised: 28-06-2024

Accepted: 02-12-2024

Keywords:

Artificial Intelligence;

Eco-Literacy;

Integrating Digital Literacy;

Islamic Education;

Religious Moderation.



© 2024 Roikhatul Jannah, Saeful Bahri, Ahmad Hufron

This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).

INTRODUCTION

The context of this study lies at the intersection of two transformative phenomena that increasingly shape Indonesia's contemporary socio-religious landscape. The first is the widespread penetration of digital technology, particularly among younger generations, which has fundamentally reshaped how individuals interact, communicate, and construct their social identities (Tika Dewi Amelia & Rania Balqis, 2023). University students, as the most digitally immersed demographic group, now regard cyberspace as a primary arena for self-expression, public discourse, and social engagement. The second phenomenon concerns the strengthening of religious moderation discourse and policy as a national strategic agenda. Within Indonesia's highly multicultural society, religious moderation is viewed as a crucial framework for maintaining interfaith harmony, sustaining social cohesion, and countering the spread of radical and extremist ideologies (Mala & Hunaida, 2023).

Within this policy landscape, higher education institutions—particularly Islamic Higher Education Institutions (PTKI)—are positioned as key agents in disseminating and internalizing the values of religious moderation (Afwadzi & Miski, 2021). Institutions such as the Institut Agama

Islam Bakti Negara (IAIBN) Tegal serve as social laboratories where moderate, tolerant, and balanced Islamic values are systematically embedded through curriculum design, student activities, and dedicated programs such as the “Rumah Moderasi.” These initiatives aim to cultivate graduates who are not only academically competent but also embody Islam as rahmatan lil ‘alamin (a mercy to all creation) (Muhsin et al., 2024).

Despite these efforts, a fundamental challenge remains largely unaddressed. While moderation values are taught and reinforced within structured academic environments, the real test of internalization occurs in cyberspace—a setting that is inherently fluid, unregulated, and governed by distinct psychological dynamics. The conceptual and practical gap between classroom-based pedagogy and online interaction represents a critical concern. The success of religious moderation programs cannot be assessed solely through students’ cognitive understanding on campus; rather, it must be evaluated based on the extent to which these values persist, manifest, and guide behavior within digital environments that often foster psychological disinhibition—conditions in which individuals feel freer to act outside social norms (Niklová & Hanesová, 2024). This study aims to bridge this gap by analyzing how the internalized values of religious moderation among IAIBN Tegal students shape their social morality when confronted with the complex realities of cyberspace.

The conceptual foundation of religious moderation (wasathiyah Islam) in Indonesia is grounded in Islamic teachings and local wisdom (Sutrisna, 2021). It emphasizes balance (tawazun), justice (i’tidāl), and the avoidance of extremism in both conservative and liberal forms (Asnita et al., 2024). Its theological basis is found in the Qur’an, Surah al-Baqarah verse 143, which describes Muslims as ummatan wasatan—a middle nation entrusted with maintaining equilibrium and serving as moral witnesses. The practical expression of religious moderation involves tolerance (tasāmuḥ), harmony between reason and revelation, and constructive dialogue that reduces prejudice. In the Indonesian context, this framework functions as both a theological imperative and an institutional strategy for strengthening multicultural citizenship (fiqh al-muwāṭanah).

The internalization of moderation values is a psychological process through which external norms become integrated into an individual’s belief system and personality (Widiastuti, 2023). This process involves cognitive (understanding), affective (appreciation), and conative (action) dimensions (Alhamid et al., 2024). Qualitative studies outline four stages of internalization: rejection, comprehension, trial practice, and integration—where values become embedded within one’s identity (Vongkulluksn et al., 2017). Internal factors such as motivation and external factors such as social environment influence transitions across these stages (Lestari et al., 2022). Social morality—understood as prosocial behaviors including fairness, respect, and kindness—represents the behavioral expression of moderation values in cyberspace. Empathy plays a central role in shaping these behaviors, comprising both cognitive empathy (perspective-taking) and affective empathy (emotional resonance) (Luis et al., 2023; Yalçın & DiPaola, 2019). While empathy promotes altruism and reduces antisocial behavior (D’Osia, 2024), recent evidence suggests a decline in empathy among university students, potentially weakening their social morality online.

To understand the complexities of online behavior, this study draws on John Suler’s Online Disinhibition Effect, which explains why individuals behave more freely or impulsively in digital environments. Six factors—dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination, and minimization of authority—may lead either to benign disinhibition (positive openness) or toxic disinhibition (negative behaviors such as hostility or hate speech). As an ambivalent environment, cyberspace can either reinforce or undermine social morality depending on how individuals navigate these psychological dynamics.

Recent empirical and conceptual work has begun to examine the relationship between higher-education efforts to promote religious moderation and students’ behavior in broader social contexts (Mukhibat et al., 2024; Shofiyuddin et al., 2024). Parallel research has explored how digital environments change social interaction and psychological processes—showing both opportunities for constructive engagement and risks of dehumanizing, disinhibited behavior online (Mustapha, et

al., 2022; Hagen et al., 2024). At the same time, studies on value internalization emphasize that cognitive understanding does not always translate into stable behavioral dispositions without experiential practice and reflective skills (Zittoun & Gillespie, 2015). Previous studies on religious moderation within PTKI environments have shown that institutional programs and curricular interventions generally strengthen students' conceptual understanding of moderation values, although the depth of internalization varies across contexts (Isnaniah & Islahuddin, 2022; Marhamah et al., 2024; Marjani et al., 2023). Research in the field of digital behavior has further demonstrated that online interactions are strongly influenced by psychological factors such as anonymity, invisibility, and reduced social cues, which can either facilitate openness or encourage impulsive and hostile communication (Niklová & Hanesová, 2024; Suler, 2004). Studies on empathy and social behavior highlight that digital settings often diminish affective resonance and perspective-taking, leading to inconsistent expressions of prosocial conduct among university students (Luis et al., 2023). Meanwhile, value-internalization research indicates that individuals commonly progress through stages ranging from basic comprehension to deeper integration, with various internal and external factors shaping the stability of these values in daily life (Vongkulluksn et al., 2017). Collectively, these studies reveal that religious moderation education, digital psychological dynamics, empathy levels, and value-internalization processes each play significant roles in shaping students' attitudes and behaviors in online and offline environments.

In this context, the study is centered around the following key research question: How does the internalization of religious moderation values at IAIBN Tegal influence students' social morality in cyberspace? To delve into this primary inquiry, the study presents three subsidiary questions: (1) How do IAIBN Tegal students comprehend and internalize the values of religious moderation introduced through formal institutional programs? (2) In what ways are students' social morality and prosocial behaviors demonstrated across various online platforms? (3) What psychological characteristics of cyberspace significantly influence these behaviors, and how do they interact with the internalized values of religious moderation?

METHODS

This study adopted a qualitative case-study design to provide a comprehensive and context-sensitive understanding of how religious moderation values are internalized and expressed within digital environments by students at IAIBN Tegal (Adu-Marfo et al., 2024). The case-study approach was selected because it allows the researcher to examine complex social phenomena within their real-life setting, particularly the interplay between institutional programs, personal value formation, and online behavioral practices. IAIBN Tegal was chosen through purposive sampling due to its active implementation of religious moderation initiatives mandated by the Ministry of Religious Affairs, making it a relevant and information-rich site for exploring the relationship between value internalization and digital interaction. Participants included students engaged in campus moderation programs, along with lecturers and program administrators who were directly involved in designing or implementing these initiatives.

Data were collected through three complementary techniques to ensure depth and triangulation. First, semi-structured in-depth interviews were conducted using a flexible interview guide that enabled participants to describe their understanding, experiences, and behavioral tendencies related to religious moderation and online engagement (Mahat-Shamir et al., 2019). Second, digital participatory observation (digital ethnography) was employed to document naturally occurring interactions on public social-media platforms, including comment threads, discussion forums, and student-managed online communities (Jensen et al., 2022). This method allowed the researcher to observe authentic behavioral expressions influenced by digital dynamics such as anonymity and asynchronicity. Third, document analysis was carried out on institutional records, curriculum documents, program guidelines, and public communication materials to understand how moderation values are formally introduced and reinforced at the institutional level. Throughout the

data collection process, the researcher used structured tools—such as interview protocols, field-note templates, and digital observation checklists—to maintain consistency and analytic rigor.

The analytical process followed the Miles and Huberman interactive model (Asipi et al., 2022), consisting of data reduction, data display, and conclusion drawing. All interview transcripts, digital field notes, and institutional documents were manually coded, with initial open coding used to identify preliminary patterns, followed by axial coding to refine categories, and selective coding to establish core themes. The analysis was guided by three theoretical frameworks that shaped both the coding structure and the interpretive lens. Value Internalization Theory provided a conceptual basis for categorizing participant responses into stages of value adoption—from basic comprehension to deeper integration. John Suler’s Online Disinhibition Effect offered a psychological lens to interpret variations in students’ online behavior through mechanisms such as dissociative anonymity, invisibility, solipsistic introjection, and minimization of authority (Suler J., 2004; Yurchenko N. I., 2024). Additionally, Empathy Theory, particularly the distinction between cognitive and affective empathy, informed the interpretation of prosocial and antisocial tendencies emerging in digital interactions (Marin-Lopez et al., 2019; Chang et al., 2021). The integration of these theories ensured that both cognitive and contextual factors were considered during analysis, resulting in a holistic interpretation of the qualitative data.

RESULTS AND DISCUSSION

Institutional Programs Supporting the Internalization of Moderation Values

The findings of this study reveal that the Institut Agama Islam Bakti Negara (IAIBN) Tegal, as part of the broader ecosystem of Islamic Higher Education Institutions (PTKI), has implemented a comprehensive set of structured and formal strategies to internalize religious moderation values among students. These institutional efforts align closely with the Ministry of Religious Affairs’ national directives and take three major forms. The first is the establishment of the Religious Moderation House, which serves as a center for study, training, mentoring, and dissemination of moderation discourse on campus. This initiative mirrors similar programs in other PTKIN, including IAIN Ponorogo, and functions as both a preventive mechanism against radical ideological infiltration and a platform for promoting tolerant and inclusive Islamic perspectives.

Second, the integration of moderation values into the curriculum is executed through key courses such as Islamic Studies, Pancasila and Civic Education, and Fiqh. These courses introduce essential concepts of *wasathiyah*, interreligious tolerance, and national commitment, ensuring that moderation is embedded intellectually within students’ academic experience. Third, IAIBN Tegal provides extensive extracurricular exposure through seminars, public lectures, workshops, student forums, and dialogue sessions that highlight themes of citizenship, pluralism, and intergroup harmony. These activities reinforce formal learning and create opportunities for experiential engagement with moderation values.

Beyond the structured programs already implemented by IAIBN Tegal, interviews revealed that repeated exposure to moderation discourses—both in classrooms and informal student circles—shaped students’ perceptions and attitudes in subtle ways. One participant explained how moderation themes frequently emerged during casual discussions with peers, *“My friends in the organization often discuss issues of tolerance and radicalism. Even though there are no classes, these discussions continue and help shape my perspective on differences”* (Informant 11). Another student described how discussions initiated by lecturers outside class hours deepened their understanding, *“Sometimes, after class, the lecturer would discuss moderation and social conditions. From that, I understood that moderation is not just a theory but must be reflected in our behavior”* (Informant 14). These insights illustrate how moderation education extends beyond formal structures, shaping students through repeated social interactions and informal learning channels.

Interviews with students indicate that these institutional mechanisms have been effective in cultivating a cognitive-level understanding of religious moderation. Many students were able to

articulate moderation as tolerance, rejection of extremism, respect for diversity, and commitment to balanced thinking. They also demonstrated awareness of the normative and theological foundations underlying these values, showing that cognitive comprehension has generally been achieved across the student population.

The students' ability to articulate moderation values aligns strongly with the definition and indicators of *moderasi beragama* formulated by Lukman Hakim Saifuddin, which include national commitment, tolerance, anti-violence, and accommodation of local culture. The findings demonstrate that IAIBN Tegal's programs succeeded in transmitting these principles at the conceptual level, echoing prior research by Afwadzi & Miski, (2021), Umar et al. (2024), and Danisworo et al. (2024), which show that PTKI-based moderation programs generally enhance students' intellectual understanding of moderation. When compared to Lukman Hakim Saifuddin's four indicators— national commitment, tolerance, anti-violence, and accommodating local culture—the findings show that students primarily internalize the tolerance and anti-violence dimensions at the conceptual level. Their ability to articulate moderation as *menghindari ekstremisme, menghormati perbedaan, dan bersikap adil* reflects alignment with LHS's framework.

Stages of Value Internalization Among Students

Despite strong cognitive comprehension, many students' understanding of religious moderation remains located at the normative–conceptual level rather than developing into a deeply internalized personal conviction. Analysis of student responses shows that IAIBN Tegal's programs have generally succeeded in shifting students from the neglect stage to the comprehension stage of value internalization. A smaller proportion—typically those actively involved in organizations or exposed to diverse social environments—display indicators of progressing to the practice–trial stage. These students reported applying tolerance in daily interactions, being more cautious in interpersonal communication, and avoiding confrontational attitudes on campus.

Only a limited number of participants appeared to reach the integration stage, wherein moderation values become intrinsically embedded within one's identity and consistently guide behavior across contexts. This trend aligns with studies from other state Islamic universities, which highlight that top-down, textually oriented approaches often produce shallow internalization—more conceptual than transformative in terms of character development.

Several participants acknowledged that despite understanding moderation values conceptually, real-life application—especially online—often presented personal challenges. One student voiced this tension clearly, *“I know I should be patient and not emotional, but it's very difficult to restrain myself when comments offend my religion”* (Informant 12). Another participant articulated a moral struggle between ideal values and emotional impulses, *“After online debates, I sometimes feel guilty for my comments, realizing they were not appropriate. At the time, however, it felt justified because I was defending my beliefs”* (Informant 13). These expressions reveal persistent inconsistencies between ideal values and lived digital behavior, highlighting an internal moral ambivalence intensified by online contexts.

While cognitive comprehension appears strong, the progression toward deeper internalization remains limited. According to Value Internalization Theory (Buckley, 2016), values are not fully integrated until they penetrate the affective and conative domains. The findings indicate that most students remain at the conceptual level, supporting studies by Muslim (2024) and Anshori & Sahara (2020), which suggest that internalization requires consistent experiential reinforcement rather than reliance on textual instruction alone. The gap between knowing and doing—especially in emotionally charged online contexts—illustrates the fragility of value internalization in the absence of metacognitive and affective training.

Divergent Expressions of Moderation in Digital Environments

Digital observations and in-depth interviews reveal a more complex and sometimes paradoxical pattern in how students express moderation values online. Their behavior in cyberspace

tends to diverge significantly from their offline conduct, indicating situational inconsistency in embodying moderation. The findings show two contrasting behavior patterns that correspond to John Suler’s concepts of benign disinhibition and toxic disinhibition.

Some students exhibited benign disinhibition, using the digital space for constructive self-expression. Students who were usually reserved in face-to-face settings became more vocal online, engaging in deep discussions on sensitive religious issues, sharing pluralistic perspectives, or mobilizing digital initiatives such as fundraising for social causes. These behaviors demonstrate the potential of digital environments to amplify openness, empathy, and prosocial engagement when supported by a positive value foundation.

In contrast, toxic disinhibition appeared prominently as well. During online debates involving politics, religion, or identity-sensitive issues, some students used confrontational language, engaged in ad hominem attacks, spread unverified information, or posted derogatory content targeting specific groups. Such behaviors contradict moderation values including tolerance, balance, and constructive dialogue. These patterns demonstrate how the psychological affordances of digital environments can undermine values that students claim to embrace.

Students frequently described emotional triggers—such as provocative posts, misinformation, or identity-related issues—as key factors that pushed them toward harsh or impulsive responses. As one participant stated, *“When I read posts that mislead or insult Islam, I immediately feel angry. Sometimes, I respond without thinking twice”* (Informant 15). Another student shared similar experiences, *“When I see extreme political comments, I automatically get provoked. I can’t help but join the debate”* (Informant 17). Such testimonies highlight how digital platforms amplify emotional reactivity, creating conditions where moderation values become temporarily overshadowed by affect-driven responses.

Students’ paradoxical behavior—moderate offline but confrontational online—reflects the psychological mechanisms elaborated in Suler’s Online Disinhibition Effect (2004). Dissociative anonymity and invisibility lower social inhibitions, reduce empathy, and create a sense of consequence-free expression. As documented by Niklová & Hanesová (2024), such digital affordances often distort interpersonal dynamics, intensifying emotional reactivity and diminishing moral accountability. These results confirm that online contexts can temporarily override internalized values, revealing that moral behavior is highly situational.

Table 1. Online Disinhibition Factors and Their Manifestations Among IAIBN Tegal Students

Online Disinhibition (Suler, 2004)	Manifestations in Students’ Online Behavior (Field Evidence)	Implications for Social Morality & Alignment with Moderation Values
Dissociative Anonymity (You Don’t Know Me)	Using fake or anonymous accounts to post harsh criticism of campus policies or engage in cyberbullying toward peers.	Direct Contradiction: Removes moral accountability by separating the “online self” from the “offline self,” conflicting with Islamic principles of integrity and ethical conduct.
Invisibility (You Can’t See Me)	A typically shy female student becomes an active admin in an online group promoting interfaith dialogue and offering emotional support to peers in distress.	Potential Alignment (Benign): Enables marginalized voices to participate. The absence of physical judgment fosters openness and empathy, supporting inclusivity within moderation values.

Asynchronicity (See You Later)	Posting provocative or hateful comments on sensitive news topics and immediately logging out, avoiding clarification or continued dialogue.	Direct Contradiction: “Hit-and-run” communication undermines patient, responsible, and constructive dialogue—core elements of moderate conflict resolution.
Solipsistic Introjection (It’s All in My Head)	Misinterpreting a friend’s text (from a different political orientation) as a hostile personal attack, escalating into conflict and online estrangement.	Direct Contradiction: Obstructs cognitive empathy by prompting individuals to project their assumptions onto others, inhibiting fair and balanced (<i>i’tidāl</i>) perspective-taking.
Dissociative Imagination (It’s Just a Game)	Participating in “meme wars” involving racial or religious stereotypes, justified as harmless entertainment (“for the lulz”).	Direct Contradiction: Creates moral detachment by treating online behavior as a game, despite real-world consequences that damage social harmony.
Minimization of Authority (We’re Equals)	Publicly and critically challenging a well-known religious figure in social media comment sections, supported by data and references—behavior unlikely to occur in face-to-face settings.	Ambivalent: Can reflect positive critical thinking aligned with balancing reason and revelation. However, when lacking respect and proper decorum (<i>adab</i>), it risks undermining social order and civility.

The findings demonstrate that the internalization of religious moderation values is critically needed to regulate the Maladaptive/Toxic Disinhibition that flourishes in the digital environment. The online space is not a moral vacuum; without a strong ethical anchor—specifically one rooted in social harmony, accountability, and balanced dialogue—students tend to regress to impulsive, divisive, and morally dissociated behaviors. Therefore, the institution's efforts must focus on integrating moderation values as a conscious mechanism of cognitive-ethical control to manage the inherent freedom of online interaction, ensuring that the student's digital *akhlāq* reflects their offline commitment to *wasatiyyah*.

Online Disinhibition as a Psychological Mechanism Shaping Student Behavior

Analysis using the six dimensions of the Online Disinhibition Effect—dissociative anonymity, invisibility, asynchronicity, solipsistic introjection, dissociative imagination, and minimization of authority—shows how specific digital conditions influence students’ moral behavior. Dissociative anonymity enables students to detach their online persona from their offline identity, reducing feelings of accountability. Invisibility, by removing nonverbal cues, weakens affective empathy and heightens emotional distance. Asynchronicity allows impulsive “hit-and-run” communication, where students post harsh comments and log out without facing immediate social repercussions (Table 1).

Solipsistic introjection leads students to project personal assumptions or biases onto written text, increasing the likelihood of misunderstanding. Dissociative imagination encourages them to treat online interactions as a “game” separate from real-world ethics. Minimization of authority

gives students a heightened sense of equality with figures they would typically treat with greater respect offline, leading to bolder expressions of criticism or disagreement.

Students were explicit about how anonymity emboldened them to express thoughts they would otherwise suppress, *“When I use an anonymous account, I feel free. I don't care what people say because they don't know who I am”* (Informant 16). Others acknowledged that anonymity encouraged behavior inconsistent with their values, *“The second account made me brave enough to speak up. That's not my style when meeting people face to face”* (Informant 18). These statements reinforce the centrality of online disinhibition in shaping digital interactions, confirming that anonymity significantly reduces self-regulation.

The erosion of empathy in cyberspace plays a decisive role in weakening students' ability to enact moderation values. This aligns with findings by Luis et al. (2023) and Yalçın & DiPaola, (2019), who argue that empathy is foundational to prosocial behavior and moral judgment. When empathy declines due to the absence of nonverbal cues and emotional immediacy, abstract values such as “respecting differences” become difficult to manifest. Conversely, when empathy is activated, students demonstrate more constructive and dialogical behavior.

Empathy Fluctuations and Their Impact on Social Morality Online

At the core of these behavioral inconsistencies lies a noticeable fluctuation in empathy. Students reported that digital interactions make it more difficult to perceive the emotions of others, reducing affective resonance and weakening moral restraint. The decline of empathy contributes to more impulsive or aggressive online conduct, as previously documented by Luis et al. (2023) and Yalçın & DiPaola (2019). The findings show that empathy plays a central role in mediating moderation values in cyberspace: when empathy diminishes, the ability to uphold moderation weakens; when empathy is activated—such as in supportive or prosocial digital communities—moderate behavior becomes more pronounced.

The findings highlight a significant gap in current religious moderation education: while cognitive transmission is strong, students lack the metacognitive and emotional competencies needed to sustain moderation values in cyberspace. As your manuscript rightly notes, students are not trained to recognize how dissociative anonymity can distort moral identity or how asynchronicity can trigger impulsive responses. The absence of Digital Moderation Literacy—a framework that strengthens self-regulation, empathy, and awareness of digital psychological dynamics—emerges as a crucial weakness.

As suggested by Grignoli, et al. (2022), empathy and prosocial behavior require active practice in contexts that resemble real-world dilemmas. Thus, moderation education must shift beyond theological instruction toward experiential methods, reflective digital simulations, and structured online engagements that cultivate emotional resilience and ethical digital citizenship.

CONCLUSION

The study concludes that while the internalization of religious moderation values among IAIBN Tegal students has been largely successful at the cognitive level, the translation of these values into consistent social morality in cyberspace remains complex and situational. Institutional programs such as the Religious Moderation House and curriculum integration effectively strengthen students' theoretical understanding, yet this awareness does not automatically manifest in moderate online behavior due to the powerful psychological dynamics of digital interaction. The Online Disinhibition Effect—particularly dissociative anonymity, asynchronicity, and dissociative imagination—creates conditions that weaken empathy and produce a gap between offline moral identity and online persona, enabling both moderate and toxic behaviors to emerge from the same individuals depending on context. These findings highlight that value internalization is a dynamic, context-dependent process shaped by the affordances of digital environments, underscoring the need for moderation education to shift from mere knowledge transmission toward metacognitive and moral skill-building tailored to online interaction. Accordingly, the development of a “Digital

Moderation Literacy” curriculum—featuring training in self-regulation, empathy, and recognition of disinhibition factors—is strongly recommended, supported by experiential learning methods and national policy alignment. Despite limitations related to institutional scope and the qualitative nature of the data, the study offers key insights that warrant further comparative, quantitative, and longitudinal research to advance adaptive and effective strategies for strengthening religious moderation in the digital age.

REFERENCES

- Adu-Marfo, A. O., Kwapong, O. A. T. F., Oheneba-Sakyi, Y., & Miller-Young, J. (2024). Understanding teachers’ usage of YouTube as a pedagogical tool: A qualitative case study of basic school teachers in Ghana. *E-Learning and Digital Media*, 0(0). <https://doi.org/10.1177/20427530241239393>
- Afwadzi, B., & Miski, M. (2021). Religious Moderation In Indonesian Higher Educations: Literature Review. *Ulul Albab Jurnal Studi Islam*, 22(2), 203–231. <https://doi.org/10.18860/ua.v22i2.13446>
- Alhamid, I. S., Nur, I., & Hasbullah. (2024). Internalisasi Nilai-Nilai Pendidikan Islam Dalam Pembentukan Karakter Pada Peserta Didik Di Sd Inpres 2 Wagom. *Transformasi : Jurnal Kepemimpinan & Pendidikan Islam*, 7(2), 29–56. <https://doi.org/10.47945/transformasi.v7i2.1550>
- Anshori, I., & Sahara, E. E. (2020). Internationalization Of Islamic Education Institutional Characters Through The English Acculturation Strategy. *Humanities & Social Sciences Reviews*, 8(5), 140–157. <https://doi.org/10.18510/hssr.2020.8514>
- Asipi, L. S., Rosalina, U., & Nopiyadi, D. (2022). The analysis of reading Habits using Miles and Huberman Interactive Model to empower students’ literacy at IPB Cirebon. *International Journal of Education and Humanities*, 2(3), 117–125. <https://doi.org/10.58557/ijeh.v2i3.98>
- Asnita, D., Kurnia, R. D., Pertiwi, A., Hasan, Z., Muhtadi, R., & Mui, P. (2024). *Moderasi Beragama dalam Konteks Studi Syariah Tantangan dan Pemahaman Yang Beragama. September*. Google
- Buckley, P. (2016). Internalization theory. In *Palgrave Macmillan UK eBooks* (pp. 1–6). https://doi.org/10.1057/978-1-349-94848-2_423-1
- Chang, S. A., Tillem, S., Benson-Williams, C., & Baskin-Sommers, A. (2021). Cognitive empathy in subtypes of antisocial individuals. *Frontiers in Psychiatry*, 12, 677975. <https://doi.org/10.3389/fpsy.2021.677975>
- D’Osia, A. V. (2024). *Promoting Well-Being in the Workplace : The Impact of Empathy , Prosocial Behaviors , and Prospective Thoughts in the Positive Psychology Framework Armelle Vallette d’ Osia*. <https://doi.org/https://doi.org/10.35662/unine-thesis-3136>
- Danisworo, T. G., Rahmatunissa, M., & Paskarina, C. (2024). Religious Moderation in Indonesia: Navigating Conflict and Pluralism through Agonistic Democracy. *MUHARRIK: Jurnal Dakwah Dan Sosial*, 7(2), 183–203. <https://doi.org/10.37680/muharrik.v7i2.5625>
- Grignoli, N., Filippini, C., & Petrocchi, S. (2022). Eliciting empathetic drives to prosocial behavior during stressful events. *Frontiers in Psychology*, 13, 963544. <https://doi.org/10.3389/fpsyg.2022.963544>
- Hagen, S., Burat, A. ., & Kahale, H. . (2024). Psychological Dynamics of Social Interaction in Face-to-Face and Digital Communication. *Journal Social Civilecial*, 2(1), 11–21. <https://doi.org/10.71435/610485>
- Isnaniah, S., & Islahuddin, I. (2022). The Application Of Religious Moderation Concept In Indonesian Language Teaching: A Case Study In Islamic Schools And Islamic Universities (Perguruan Tinggi Keagamaan Islam/Ptki). *Penamas*, 35(2), 213-222. <https://doi.org/10.31330/penamas.v35i2.629>
- Jensen, L. X., Bearman, M., Boud, D., & Konradsen, F. (2022). Digital ethnography in higher education teaching and learning—a methodological review. *Higher Education*, 84(5), 1143–

1162. <https://doi.org/10.1007/s10734-022-00838-4>

- Lestari, A. F., Nurdin, N., & Hamka, H. (2022). *Proceeding of International Conference on Islamic and Interdisciplinary Studies (ICIIS), 2022 Factors Influencing Value Internalization On Students*. 668–672. [Google](#)
- Luis, E. O., Martínez, M., Akriovou, K., Scalzo, G., Aoiz, M., & Orón Semper, J. V. (2023). The role of empathy in shared intentionality: Contributions from Inter-Processual Self theory. *Frontiers in Psychology, 14*(March), 1–17. <https://doi.org/10.3389/fpsyg.2023.1079950>
- Mahat-Shamir, M., Neimeyer, R. A., & Picho-Prelorntzos, S. (2019). Designing in-depth semi-structured interviews for revealing meaning reconstruction after loss. *Death Studies, 45*(2), 83–90. <https://doi.org/10.1080/07481187.2019.1617388>
- Mala, A., & Hunaida, W. L. (2023). Exploring the Role of Religious Moderation in Islamic Education: A comprehensive analysis of its unifying potential and practical applications. *Jurnal Pendidikan Agama Islam (Journal of Islamic Education Studies), 11*(2), 173–196. <https://doi.org/10.15642/jpai.2023.11.2.173-196>
- Marhamah, M., Rohana, S., Putra, R. S., Khatami, M., & Hamarita, M. (2024). Religious Moderation Literacy Index of PTKIN Students in Aceh: Analysis of Islamic Knowledge and Understanding. *Analisis Jurnal Studi Keislaman, 24*(2), 285–312. <https://doi.org/10.24042/ajsk.v24i2.23355>
- Marín-López, I., Zych, I., Monks, C.P., Ortega-Ruiz, R. (2019). Empathy, Morality and Social and Emotional Competencies in Interpersonal Interactions Online. In: Coetzee, M. (eds) *Thriving in Digital Workspaces*. Springer, Cham. https://doi.org/10.1007/978-3-030-24463-7_11
- Marjani, G. I. (2023). Overcoming Theological Dilemmas: Fostering Religious Moderation through the Resolution of Faith and Rationality. *Religious Jurnal Studi Agama-Agama Dan Lintas Budaya, 7*(2), 115–128. <https://doi.org/10.15575/rjsalb.v7i2.28311>
- Muhsin, M., Kususiyannah, A., & Maksum, M. (2024). Religious Moderation in Indonesian Islamic Universities: Policy Implementation and Identity Formation at IAIN Ponorogo. *Ascarya: Journal of Islamic Science, Culture and Social Studies, 4*(2), 54–66. <https://doi.org/10.53754/iscs.v4i2.713>
- Mukhibat, M., Effendi, M., Setyawan, W. H., & Sutoyo, M. (2024). Development and evaluation of religious moderation education curriculum at higher education in Indonesia. *Cogent Education, 11*(1). <https://doi.org/10.1080/2331186X.2024.2302308>
- Muslim, M. (2024). Internalising digital technology in Islamic education. *SCAFFOLDING Jurnal Pendidikan Islam Dan Multikulturalisme, 6*(3). <https://doi.org/10.37680/scaffolding.v6i3.6309>
- Mustapha, I., Khan, N., Qureshi, M.I. (2022). Is Technology Affecting the Way Our Minds Operate? Digital Psychology of Users in the Era of Digitalization. In: Ismail, A., Mohd Daril, M.A., Öchsner, A. (eds) *Advanced Transdisciplinary Engineering and Technology. Advanced Structured Materials*, vol 174. Springer, Cham. https://doi.org/10.1007/978-3-031-01488-8_8
- Niklová, M., & Hanesová, D. (2024). Religious Education as a Platform for Pupils' Social Development and Prevention of Internet Addiction: The Case of Slovakia. *Religions, 15*(5). <https://doi.org/10.3390/rel15050585>
- Shofiyuddin, A., Khoiriyah, S., & Rangga Sa'adillah. (2024). Building Tolerance And Balance: A Systematic Literature Review On Religious Moderation Among Students In Higher Education. *Journal of Islamic Civilization, 5*(2), 123–137. <https://doi.org/10.33086/jic.v5i2.5305>
- Suler, J. (2004). The Online Disinhibition Effect. *CyberPsychology & Behavior, 7*(3). <https://doi.org/10.1089/1094931041291295>
- Sutrisna, S. (2021). Local Wisdom as the Basis for Religious Moderation in Pluralistic Indonesian Society to Realize Islamic Values Rahmatan lil 'Alamin. *Millati: Journal of Islamic Studies and Humanities, 6*(2), 243–256. <https://doi.org/10.18326/mlt.v6i2.6581>
- Tika Dewi Amelia, L., & Rania Balqis, N. (2023). Changes in Communication Patterns in the

- Digital Age. *ARRUS Journal of Social Sciences and Humanities*, 3(4), 544–556. <https://doi.org/10.35877/soshum1992>
- Umar, M., Pangalila, T., & Ismail, F. (2024). Peace-Loving and Anti-Violence in Indonesian Educational Institutions. In *Character Education through Peace* (pp. 601–611). <https://doi.org/10.4324/9781003352471-74>
- Vongkulluksn, V. W., Xie, K., & Bowman, M. A. (2017). The role of value on teachers' internalization of external barriers and externalization of personal beliefs for classroom technology integration. *Computers & Education*, 118, 70–81. <https://doi.org/10.1016/j.compedu.2017.11.009>
- Widiastuti, N. (2023). Metode Pembelajaran Dalam Upaya Internalisasi Nilai-Nilai Keislaman. *Al Fatih*, 1(1), 1–8. [Google](#)
- Yalçın, Ö. N., & DiPaola, S. (2019). Modeling empathy: building a link between affective and cognitive processes. *Artificial Intelligence Review*, 53(4), 2983–3006. <https://doi.org/10.1007/s10462-019-09753-0>
- Yurchenko N.I. (2024). Обзор статьи John Suler «The Online Disinhibition Effect». *New Psychological Research*, 4(1). https://doi.org/10.51217/npsyresearch_2024_04_01_09
- Zittoun, T., & Gillespie, A. (2015). Internalization: How culture becomes mind. *Culture & Psychology*, 21(4), 477-491. <https://doi.org/10.1177/1354067X15615809>