Reconstructing Islamic Moral Education in The Age of Artificial Intelligence: An Integration of Taqwa, Qalb, and Artificial Moral Reasoning

Robiah Adawiyah
Universitas Islam Negeri K.H. Abdurrahman Wahid Pekalongan, Indonesia
Email: siapamirat@gmail.com

Abstract

The rapid development of artificial intelligence (AI) has transformed educational practices, knowledge production, and moral decision-making, raising fundamental questions about the role of moral education in shaping responsible human agency. In Islamic education, these developments highlight the need to reconsider the foundations of moral formation beyond technological literacy and cognitive competence. This study aims to reconstruct an Islamic moral education framework by integrating the concepts of *taqwa*, *qalb*, and artificial moral reasoning to address emerging ethical challenges in the age of AI. Employing a qualitative library research method, this study adopts a philosophical and conceptual approach through critical analysis of classical Islamic scholarship, particularly the moral thought of Al-Ghazali, alongside contemporary literature on AI ethics, moral cognition, and educational philosophy. The findings suggest that *taqwa* can be understood as a foundational moral orientation that guides ethical judgment and responsible action, while *qalb* represents the spiritual and moral center of human consciousness in Islamic thought. Artificial moral reasoning, meanwhile, provides a contemporary framework for examining how AI-mediated environments influence ethical reflection and decision-making. Integrating these concepts offers a framework in which technological competence is complemented by spiritual awareness, moral discernment, and ethical responsibility. This study concludes that Islamic moral education in the age of AI should not merely focus on the effective use of technology but also cultivate taqwa, strengthen moral consciousness, and preserve human agency. The proposed framework contributes to contemporary discussions on Islamic education by positioning spiritual and ethical formation as essential foundations for navigating technological transformation.

Keywords: artificial intelligence, islamic moral education, taqwa, qalb, artificial moral reasoning

Introduction

Artificial intelligence (AI) is no longer a distant technology in education. Students now draft essays, summarize readings, translate texts, and ask for advice through generative systems that answer in seconds, while teachers use similar tools to plan lessons and assess work (Kasneci et.al., 2023). Institutional responses have mostly concentrated on two matters, namely building technological literacy and preventing academic dishonesty. Both are necessary, yet neither touches a deeper question: what kind of person is being formed when moral reflection, judgment, and even the sense of responsibility are routinely shared with a machine?

This question matters because AI systems are not morally neutral. Algorithms embed choices about what counts as relevant, fair, or harmful, and those choices often remain opaque to the people affected by them (Mittelstadt et al., 2016). Vallor warns that habitual reliance on automated systems can lead to moral deskilling, a gradual weakening of the capacities for attention, deliberation, and care on which virtue depends (Vallor, 2015). Guidelines on AI ethics have multiplied, but they converge on a handful of principles and are addressed mainly to developers and institutions, not to the moral formation of learners (Jobin, et.al., 2019).

Islamic education has the conceptual resources to take up this question. Its classical aim is the formation of a good human being, not only the transmission of knowledge, a point Al-Attas (1980) expresses through the concept of *ta'dib* (Khobir et al., 2021) and Halstead through his account of the Islamic concept of education. Yet part of the current discussion on Islamic education and AI tends to revolve around using AI to deliver religious content or to equip students with digital skills. Work that engages Islamic ethics directly, such as Elmahjub's proposal for pluralist ethical benchmarking, focuses mainly on how AI systems should be evaluated and less on how learners should be formed (Elmahjub, 2023).

This article tries to fill that gap by bringing together three concepts. *Taqwa*, the God-conscious moral vigilance at the center of Qur'anic ethics, supplies the orientation of moral action. *Qalb*, the heart as Al-Ghazali understood it, supplies the inner seat of moral perception and character. Artificial moral reasoning supplies the contemporary setting, namely the ways machines simulate moral judgment and the ways AI-mediated environments shape human reflection. The guiding question is how these three concepts can be integrated into a framework for Islamic moral education in the age of AI. The article argues that technological competence has to be anchored in spiritual awareness and moral discernment; otherwise education risks producing efficient users who are morally passive.

Methods

This study uses a qualitative library research design with a philosophical and conceptual approach. It does not test hypotheses with field data. Instead, it analyzes concepts and arguments in existing texts in order to build a theoretical framework. The sources fall into three groups. The first consists of primary Islamic sources: the Qur'an, the hadith collections, and Al-Ghazali's *Ihya' 'Ulum al-Din*, especially Book 21 on the wonders of the heart, Book 22 on disciplining the soul, and Book 38 on vigilance and self-examination. Al-Ghazali was chosen because his psychology of the heart and his theory of character remain among the most systematic classical treatments of moral formation. The second group comprises scholarship on Islamic educational philosophy and Qur'anic ethics. The third covers contemporary literature on AI ethics, machine ethics, moral psychology, and AI in education. For the second and third groups, I prioritized books from academic presses and articles in peer-reviewed international journals, searched through Google Scholar and publisher databases with keywords such as *Islamic moral education*, *taqwa*, *qalb*, *machine ethics*, *moral deskilling*, and *AI in education*.

The analysis proceeded in four steps: (1) clarifying each concept through close reading of primary and secondary sources; (2) tracing points of convergence and tension among the three concepts; (3) building the framework from those relations; and (4) testing its coherence against likely objections, such as whether classical concepts can be applied to a technological setting without distortion. Qur'anic translations follow Abdel Haleem (2004). Because the study is conceptual, its findings are propositions for further empirical testing and not demonstrated effects.

Results

The analysis yields three findings. First, *taqwa* works as a moral orientation that makes ethical judgment accountable. Second, *qalb* is the faculty in which moral perception, motivation, and character are formed, which explains why moral education cannot be reduced to the transfer of information. Third, artificial moral reasoning is best treated as an object of

critical reflection, not as a moral authority. These findings are then combined into one framework.

Taqwa as a Moral Orientation

Fazlur Rahman (1980) describes *taqwa* as a form of inner alertness, an awareness of the consequences of one's actions that keeps a person on guard against wrongdoing. The Qur'an presents it as a capacity placed in the soul together with its tendency toward wrongdoing (Q 91:7–10), as the criterion by which people are honored (Q 49:13), and as the source of *furqan*, the ability to tell right from wrong (Q 8:29). Izutsu (2002) likewise treats *taqwa* as a key ethico-religious concept that joins reverence for God with moral conduct.

Three features of *taqwa* matter for education in the age of AI. It is anticipatory, because it asks what an act will mean before the act is done. It is discerning, because *furqan* is not mere rule-following but the ability to weigh competing claims. And it is accountable: the Qur'an says that hearing, sight, and the heart will each be questioned (Q 17:36), and it instructs believers to verify news brought by an unreliable person before acting on it (Q 49:6). Read together, these verses offer a sturdy ethic of information use (Anisa & Hasanah, 2025). A student who accepts an AI output without checking it, or who submits generated work as the fruit of his or her own effort, fails not only a rule of academic integrity but also the vigilance that *taqwa* demands. Al-Ghazali develops this vigilance as *muraqaba* (watchfulness) and *muhasaba* (self-reckoning), practices in which a person examines intentions before acting and reviews conduct afterwards. Applied to AI, these become habits before, during, and after use: asking why one turns to the tool, evaluating what it returns, and owning the consequences.

Qalb as the Moral Center

Al-Ghazali (2010)opens his book on the wonders of the heart by distinguishing four terms that the tradition often uses interchangeably: *qalb*, *ruh*, *nafs*, and *'aql*. For him, the heart in its spiritual sense is a subtle, divinely originated reality that knows and perceives, and it is this reality, not the physical organ, that carries moral and religious responsibility. He compares the heart to a mirror that reflects the truth when polished and fails to do so when covered by the rust of desire and habitual sin (Al Ghazali, 2010). The image fits the Qur'anic claim that what turns blind is not the eyes but the hearts within the chests (Q 22:46), the praise of the sound heart (*qalb salim*) on the Day of Judgment (Q 26:88–89), and the Prophetic saying that the whole body is sound when the one morsel of flesh, the heart, is sound (Al Bukhari, no. 1599).

Two implications follow. First, moral knowledge in this tradition is not merely propositional. Al-Ghazali (1995) defines character (*khuluq*) as a firmly rooted state of the soul from which actions issue easily and without prior deliberation, which is why he holds that moral education works through the discipline and habituation of the soul, not through instruction alone. A student can therefore know the right answer, even retrieve it from a chatbot in seconds, and still lack the disposition to act on it. Second, the condition of the heart affects how information is received. The same AI-generated answer may be used with humility or with vanity, to learn or to evade effort, depending on the person who uses it. Contemporary moral psychology offers a parallel: Haidt argues that moral judgments are often driven by quick intuitions, with reasoning following afterwards to justify them (Haidt, 2001). The parallel should not be pushed too far, since Al-Ghazali's heart is a spiritual and cognitive center and not simply an intuitive system, but it supports the claim that the formation of inner dispositions precedes and shapes reasoning.

Vallor's account of moral deskilling gives this point a technological edge. Virtues are cultivated through repeated practice, and technologies that remove the occasions for moral practice may erode the very skills they replace (Vallor, 2015). If an AI system drafts the apology, proposes the answer to a dilemma, and recommends the course of action, the heart has fewer chances to be exercised. The *qalb* perspective therefore suggests that the main danger of AI in moral education is less the spread of false information than the quiet outsourcing of inner effort.

Artificial Moral Reasoning and Its Limits

The term artificial moral reasoning is used here in two connected senses. In the narrow sense, it refers to attempts to build systems that take moral considerations into account when they act or recommend. Moor distinguishes ethical-impact agents, implicit ethical agents, explicit ethical agents, and full ethical agents, and he treats implicit ethical agents as the realistic target of present machine design, while explicit and full ethical agency remain far more difficult (Moor, 2006). Wallach and Allen (2009) speak of "functional morality," meaning machines able to assess some of the moral consequences of their actions, and they compare top-down (rule-based), bottom-up (learning-based), and hybrid designs. In the broad sense, the term refers to how AI-mediated environments, including recommendation systems and chatbots, frame options, suggest answers, and thereby influence the moral reflection of human users.

The literature calls for caution in both senses. Even scholars who grant artificial agents a limited form of moral agency distinguish accountability from responsibility, since a system cannot answer for its actions as a person can (Sanders, 2004). Large language models generate plausible text from statistical patterns, which is not the same as understanding what is said (Bender et.al., 2021). The values such systems reflect depend on data and design choices. The Moral Machine experiment, which gathered millions of decisions from people in many countries, found marked cultural variation in moral preferences, a reminder that there is no culturally neutral moral baseline on which to train (Awad et.al., 2018). Ethics guidelines show a similar pattern: a review of 84 documents found convergence on a few principles but substantial divergence in how they are interpreted and applied (Jobin, et.al., 2019). Experimental work also suggests that people tend to be averse to machines making moral decisions, which points to a lingering intuition that moral judgment is a human responsibility (Bigman, 2018).

From an Islamic perspective, the consequence is straightforward. As I read the tradition, moral accountability (*taklif*) rests on a person who has reason, intention, and a heart, and a system that has none of these cannot be its bearer. AI can assist deliberation but cannot replace the accountable human subject. Artificial moral reasoning remains valuable for education, but as a mirror. When students compare a chatbot's answer to an ethical dilemma with the criteria of *taqwa* and the condition of the heart, they can see the assumptions built into the answer and the ones they hold themselves.

Discussion

An Integrated Framework for Moral Education

The three concepts relate to one another in a cycle and not in a hierarchy. *Qalb* is the inner foundation: it is where moral perception and character are formed and where sincere intention begins. *Taqwa* is the orientation that carries this inner state into action as vigilant and accountable judgment. Artificial moral reasoning is the environment in which this judgment is

now exercised, and it both tests and shapes it. An attentive heart produces *taqwa*, *taqwa* disciplines the use of AI, and the encounter with AI in turn either strengthens the heart through reflective use or dulls it through uncritical dependence. Table 1 summarizes the framework.

Table 1. An integrated framework of *qalb*, *taqwa*, and artificial moral reasoning

Component	Role in the framework	Educational implication
Qalb	Inner foundation: moral perception, intention, and character	Habituation, reflection, and spiritual discipline; the teacher as moral exemplar
Taqwa	Orientation: anticipatory, discerning, and accountable action	Watchfulness before, during, and after AI use; verification (*tabayyun*) as an ethical habit
Artificial moral reasoning	Context: machine simulation of moral judgment and AI-mediated reflection	Critical analysis of AI outputs as case material; responsibility stays with the human user

Four educational implications follow. First, curricula should include structured reflection on AI use, for example short prompts that ask students to state their intention before using a tool and to review afterwards what they delegated and what they decided themselves. This is a classroom form of *muraqaba* and *muhasaba*. Second, verification should be taught as an ethical practice grounded in *tabayyun*, not only as a research skill. Third, AI-generated answers to moral questions can serve as case material: students identify the assumptions, omissions, and sources of authority behind them and compare these with Qur'anic and Ghazalian criteria. Fourth, the teacher's role needs to be reaffirmed. Al-Ghazali treats the disciplining of the soul as work done with guidance, and the tradition places the exemplary teacher (*murabbi*) at the center of moral formation, a role no system can fully assume because it rests on lived example and relationship (Hidayat & Fatmawati Nur Khasanah, 2025).

Two caveats are in order. The framework does not reject AI. AI can support learning and can even provide simulated dilemmas for moral discussion, and scholars of responsible AI rightly call for human-centered and accountable systems (Floridi et.al., 2018). Nor does the framework claim that Al-Ghazali's categories can be transposed mechanically. They are used here as analytical lenses, and applying them to digital settings requires careful contextual interpretation that should be tested through classroom research.

Conclusion

This article has argued that Islamic moral education in the age of AI needs a foundation deeper than technological competence. The analysis suggests that *qalb* provides the inner basis of moral perception and character, that *taqwa* gives this basis a vigilant and accountable orientation, and that artificial moral reasoning supplies a contemporary setting in which both are tested. Integrating the three yields a framework in which the use of AI is subordinated to moral formation and human agency is preserved: the one who asks, evaluates, and answers for the outcome remains the learner, not the machine.

The study offers Islamic educators and curriculum designers a conceptual vocabulary for responding to AI without reducing moral education to digital literacy. Its limits are plain. It is a conceptual study built on selected sources, and the framework has not yet been tested in classrooms. Future research could use design-based research or classroom action research to translate the framework into learning activities, develop instruments for assessing reflective awareness as the observable side of *taqwa*, and compare the Ghazalian approach with other Islamic ethical traditions.

References

Abdel Haleem, M. A. S., trans. *The Qur'an*. Oxford: Oxford University Press, 2004.

Al-Attas, Syed Muhammad Naquib. *The Concept of Education in Islam: A Framework for an Islamic Philosophy of Education*. Kuala Lumpur: Muslim Youth Movement of Malaysia, 1980.

Al-Bukhari, Muhammad ibn Isma'il. *Sahih al-Bukhari*.

Al-Ghazali. *Ihya' 'Ulum al-Din*. 4 vols. Beirut: Dar al-Ma'rifa, n.d.

_____. *On Disciplining the Soul and on Breaking the Two Desires: Books XXII and XXIII of the Revival of the Religious Sciences*. Translated by T. J. Winter. Cambridge: Islamic Texts Society, 1995.

_____. *The Marvels of the Heart: Book 21 of the Ihya' 'Ulum al-Din*. Translated by Walter James Skellie. Louisville, KY: Fons Vitae, 2010.

Allen, Colin, Iva Smit, and Wendell Wallach. "Artificial Morality: Top-Down, Bottom-Up, and Hybrid Approaches." *Ethics and Information Technology* 7, no. 3 (2005): 149–155. https://doi.org/10.1007/s10676-006-0004-4

Anisa, L. R. N., & Hasanah, F. N. (2025). Kode Etik dan Profesionalisme Guru untuk Meningkatkan Mutu Pendidikan. *Jurnal Pendidikan Guru Profesional*, *3*(1), 18–26.

Awad, Edmond, Sohan Dsouza, Richard Kim, Jonathan Schulz, Joseph Henrich, Azim Shariff, Jean-François Bonnefon, and Iyad Rahwan. "The Moral Machine Experiment." *Nature* 563 (2018): 59–64. https://doi.org/10.1038/s41586-018-0637-6

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. New York: ACM, 2021. https://doi.org/10.1145/3442188.3445922

Bigman, Yochanan E., and Kurt Gray. "People Are Averse to Machines Making Moral Decisions." *Cognition* 181 (2018): 21–34. https://doi.org/10.1016/j.cognition.2018.08.003

Elmahjub, Ezieddin. "Artificial Intelligence (AI) in Islamic Ethics: Towards Pluralist Ethical Benchmarking for AI." *Philosophy & Technology* 36 (2023): Article 73. https://doi.org/10.1007/s13347-023-00668-x

Floridi, Luciano, and J. W. Sanders. "On the Morality of Artificial Agents." *Minds and Machines* 14, no. 3 (2004): 349–379. https://doi.org/10.1023/B:MIND.0000035461.63578.9d

_____, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, et al. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28, no. 4 (2018): 689–707. https://doi.org/10.1007/s11023-018-9482-5

Haidt, Jonathan. "The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108, no. 4 (2001): 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Halstead, J. Mark. "An Islamic Concept of Education." *Comparative Education* 40, no. 4 (2004): 517–529. https://doi.org/10.1080/0305006042000284510

Hidayat, R. H., & Fatmawati Nur Khasanah. (2025). The Role of Qur'anic Education Centre Teachers in Integrating Technology and Children's Character Education. *Guruku: Jurnal Pendidikan Profesi Guru*, *4*(2). https://doi.org/10.19109/guruku.v4i2.27975

Izutsu, Toshihiko. *Ethico-Religious Concepts in the Qur'an*. Montreal: McGill-Queen's University Press, 2002.

Jobin, Anna, Marcello Ienca, and Effy Vayena. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1, no. 9 (2019): 389–399. https://doi.org/10.1038/s42256-019-0088-2

Kasneci, Enkelejda, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, et al. "ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education." *Learning and Individual Differences* 103 (2023): 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khobir, A., Sopiah, S., & Hasanah, F. N. (2021). A Holistic Model for Character Education in Schools (An Alternative Educational Model). *Edukasia Islamika*, *6*(2). https://doi.org/10.28918/jei.v6i2.4731

Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. "The Ethics of Algorithms: Mapping the Debate." *Big Data & Society* 3, no. 2 (2016): 1–21. https://doi.org/10.1177/2053951716679679

Moor, James H. "The Nature, Importance, and Difficulty of Machine Ethics." *IEEE Intelligent Systems* 21, no. 4 (2006): 18–21. https://doi.org/10.1109/MIS.2006.80

Muslim ibn al-Hajjaj. *Sahih Muslim*.

Rahman, Fazlur. *Major Themes of the Qur'an*. Minneapolis: Bibliotheca Islamica, 1980.

Vallor, Shannon. "Moral Deskilling and Upskilling in a New Machine Age: Reflections on the Ambiguous Future of Character." *Philosophy & Technology* 28, no. 1 (2015): 107–124. https://doi.org/10.1007/s13347-014-0156-9

Wallach, Wendell, and Colin Allen. *Moral Machines: Teaching Robots Right from Wrong*. New York: Oxford University Press, 2009.

Zawacki-Richter, Olaf, Victoria I. Marín, Melissa Bond, and Franziska Gouverneur. "Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where Are the Educators?" *International Journal of Educational Technology in Higher Education* 16 (2019): Article 39. https://doi.org/10.1186/s41239-019-0171-0