

Effectiveness of a Generative AI-Powered Conversational Bot in Teaching Hadith Sciences: Evidence from Al Qasimia University

Rassem Nouri 

Department of Principles of Religion, College of Sharia & Islamic Studies, Al Qasimia University, Sharjah, UAE

Email: rnouri@alqasimia.ac.ae

Manuscript received January 5, 2026; revised February 4, 2026; accepted May 18, 2026; published September 15, 2026

Abstract—Recent advances in Artificial Intelligence (AI) have opened new possibilities for redesigning teaching and learning in higher education, particularly through generative AI systems capable of producing context-sensitive feedback in real time. Among these, AI-powered conversational bots are emerging as promising tools to support students' understanding, engagement, and self-regulation. However, empirical research on the use of such tools in Islamic education—and specifically in Hadith Sciences—remains scarce. This study examines the effectiveness of a generative AI-powered conversational bot in an undergraduate Hadith Sciences course at Al Qasimia University, Sharjah. A course-specific chatbot was designed and constrained using authenticated Hadith terminology, the officially adopted textbook in Hadith Sciences, and instructor-approved explanations and practice items. The research adopts a convergent mixed-methods quasi-experimental design comparing a traditionally taught cohort with a cohort that additionally used the chatbot during a six-week summer term. Data were collected through structured Likert-scale questionnaires, semi-structured interviews, classroom observations, and analysis of course grades ($n = 80$). The findings indicate notable improvements in students' comprehension of Hadith classifications, self-evaluation, exam preparation, and time management among those who used the chatbot. Students reported that the bot simplified complex terminology, supported on-demand revision, and increased motivation, while also highlighting limitations related to question variety and technical accessibility. The study concludes that generative AI-powered conversational bots can function as effective supplementary tools for teaching Hadith Sciences education when implemented within clear ethical and Shari'ah-aligned frameworks and integrated carefully with curricular objectives.

Keywords—Artificial Intelligence (AI), generative artificial intelligence, chatbots, Hadith Sciences, Islamic education, higher education, digital learning

I. INTRODUCTION

In recent years, advances in Artificial Intelligence (AI) have profoundly transformed the landscape of teaching and learning. AI has moved from being a set of tools confined largely to industrial and computational domains to becoming a pervasive technology reshaping the humanities, social sciences, and education [1].

Within this broader shift, generative AI systems, capable of producing context-sensitive, human-like text and dialogue, have attracted particular attention for their potential to support open, distance, and digital education [1, 2].

Among the most prominent applications of generative AI in higher education are conversational agents or chatbots that can interact with students using natural language, provide

instant explanations, and offer personalized guidance [2].

A growing body of research suggests that AI-powered chatbots can serve as virtual tutors, question-answering assistants, and formative assessment tools, helping students to review content, clarify complex ideas, and receive immediate feedback on their understanding [2]. Nevertheless, the majority of empirical implementations and evaluations have focused on medical, technical, and experimental sciences, with relatively little attention given to Islamic studies and humanities contexts [3]. This imbalance raises important questions about how generative AI can be adapted to disciplines that are text-centered, value-laden, and methodologically sensitive, such as Hadith Sciences, and how such adaptations can respect the ethical and Shari'ah-based guidelines articulated in recent scholarly and institutional discussions [3].

Hadith Sciences occupy a central and highly specialized place within the Islamic scholarly tradition. They involve the careful study of chains of transmission (isnād), narrators' reliability, and multiple dimensions of Hadith classification, including categories based on multiplicity of transmission (mutawātir and āḥād: mashhūr, 'azīz, gharīb), status at the end of the chain (qudsī, marfū', mawqūf, maqtū'), and degrees of acceptance and rejection (ṣaḥīḥ, ḥasan, ḍa'īf). Teaching these concepts to undergraduate students is often challenging due to the density and hierarchy of the material, the limited time allocated to the course, and the diversity of students' prior knowledge and academic backgrounds. Traditional lecture-based methods may not always provide sufficient opportunities for individual practice, self-assessment, and iterative clarification of concepts that are easily confused, particularly when students lack effective digital tools for structured revision [4, 5]. These classificatory frameworks are presented systematically in the officially adopted course textbook [6].

In this context, generative AI chatbots may provide a promising complementary tool for supporting the teaching and learning of Hadith Sciences. Properly designed and supervised, a specialized chatbot can present definitions, examples, and review questions in an interactive manner, allowing students to revise at their own pace, identify their weaknesses, and receive tailored guidance [2, 4].

At the same time, the use of AI in Islamic education raises legitimate concerns related to the accuracy of content, the reliability of sources, respect for doctrinal diversity, protection of students' data, and adherence to ethical and Shari'ah-based standards [3, 7]. Recent works in Islamic jurisprudence and policy on AI emphasize the need for

authentic references, scholarly oversight, and clear pedagogical principles when deploying AI tools in religious and value-oriented education [7].

This study arises from a concrete educational initiative at Al Qasimia University in Sharjah, United Arab Emirates, which has adopted a forward-looking digital transformation agenda in higher education while maintaining a strong commitment to Islamic values and scholarly rigor [8]. Within this institutional context, a generative AI-based conversational bot was developed to support students enrolled in the Hadith Sciences course. The bot was trained on the officially adopted textbook and a curated set of authenticated Hadith sources and terminology, and was made available to students through the university's learning management system over a six-week summer term [6]. The present paper aims to evaluate the effectiveness of this chatbot in enhancing students' learning outcomes and experiences in comparison with a traditionally taught cohort, thereby contributing to the emerging literature on AI in Islamic education [3–5].

Specifically, the study seeks to answer the following overarching question: To what extent does the use of a generative AI-powered conversational bot enhance comprehension, self-evaluation, and exam preparation in an undergraduate Hadith Sciences course at Al Qasimia University? To address this question, the paper combines a descriptive and quasi-experimental design, integrating quantitative and qualitative data from student questionnaires, interviews, and course grades. By documenting both the design of the bot and its observed educational impact, the study offers an evidence-based model for integrating generative AI tools into Shari'ah-oriented curricula while respecting ethical and methodological constraints.

II. LITERATURE REVIEW

This section situates the present study at the intersection of three interrelated strands of research: (a) generative AI and chatbots in higher education, (b) the pedagogical specificity of Hadith Sciences, and (c) ethical and Shari'ah-based guidelines for the use of AI in Islamic education.

A. Generative AI and Chatbots in Higher Education

The rapid growth of AI in education has generated a rich body of conceptual and empirical research on how intelligent systems can support teaching and learning processes and enhance students' engagement and satisfaction [1].

Within this area, generative AI has emerged as a particularly influential paradigm, characterized by models capable of producing new text, explanations, and feedback rather than merely selecting from pre-programmed responses [1, 2]. These capabilities have opened possibilities for more adaptive and dialogic forms of instruction, especially in online and blended learning environments.

Recent studies in higher education report that generative AI chatbots can function as virtual tutors or learning companions, assisting students with clarifying course concepts, preparing for examinations, and engaging in self-paced revision [2, 4]. McGrath *et al.* [2] highlight that generative AI chatbots in university contexts are increasingly used to answer student questions, scaffold problem-solving, and personalize feedback in real time. Du Boulay situates

these developments within broader debates on AI in education and ethics, emphasizing both their potential to democratize access to support and the need for careful oversight [1]. Empirical work in different disciplines suggests that well-designed chatbots may improve learners' perceived understanding, motivation, and satisfaction, while also raising concerns about overreliance, content accuracy, and academic integrity [2, 4, 5]. Recent survey work in higher education also shows that students across disciplines are already integrating AI chatbots into their study routines, with notable differences by field, gender, and academic level [9].

Recent evidence also indicates that chatbots can support key learning processes that align with this study's outcomes—namely comprehension of complex concepts, self-evaluation through self-testing and metacognitive monitoring, and exam preparation via formative practice and feedback [10]. Beyond Islamic higher education, recent work in Religious Education more broadly has highlighted similar concerns about doctrinal integrity, data privacy, and the risk of over-reliance on AI tools for theological interpretation, while still recognizing their pedagogical potential when carefully governed [11]. A case study on ChatGPT also suggests that large language models can provide usable feedback to students, while still requiring expert oversight and careful instructional framing [12].

Despite these advances, most published studies focus on STEM, medical, and technical fields, often using AI systems trained on general-purpose corpora rather than highly specialized disciplinary content [2, 4]. There remains a noticeable gap with regard to the use of generative AI in text-centered, interpretive, and value-laden domains such as theology, religious studies, and Islamic sciences [3]. Mahmudulhassan *et al.* [3], for example, discuss the opportunities and challenges of AI in multicultural Islamic education and call for context-sensitive, ethically grounded implementations that respect doctrinal diversity and local educational needs. The present study responds to this gap by examining a generative AI chatbot specifically designed for an undergraduate Hadith Sciences course in a Shari'ah-oriented university setting.

B. The Pedagogical Specificity of Hadith Sciences

Hadith Sciences constitute a sophisticated and methodologically demanding field within Islamic scholarship. Students must master a wide range of technical concepts, including categories of Hadith based on multiplicity of transmission (*mutawātir* vs. *āḥād*, with its subcategories *mashhūr*, *ʿazīz*, *gharīb*), status at the end of the chain (*qudsī*, *marfūʿ*, *mawqūf*, *maqtūʿ*), and degrees of acceptance and rejection (*ṣaḥīḥ*, *ḥasan*, *ḍaʿīf*), in addition to understanding *isnād* criticism and the principles of *al-jarḥ wa-l-taʿdīl* [6]. Standard textbooks—such as *Zād al-Ṭālib min ʿUlūm al-Ḥadīth* adopted at Al Qasimia University—present these classifications systematically, but their density can be challenging for undergraduate learners [6].

From a pedagogical perspective, teaching Hadith Sciences often faces several recurring constraints: the limited number of contact hours per semester, heterogeneous student backgrounds, and the need for repeated explanation and comparison of similar-sounding categories [6].

Traditional lecture-based approaches, while indispensable

for conveying authoritative interpretations, may not provide enough space for individual practice, self-testing, and iterative feedback, especially when students are reluctant to ask questions publicly. Studies on AI in higher education suggest that digital tools can complement classroom instruction by providing additional opportunities for practice and revision outside class time [4, 5].

In the context of Hadith Sciences, a specialized chatbot can help students to revisit definitions, distinguish between closely related terms, and test themselves through formative questions aligned with the course structure.

The present study builds on this pedagogical analysis by operationalizing the core axes of the Hadith Sciences course—Hadith types by multiplicity of transmission, by end-of-chain status, and by acceptance/rejection—as the main knowledge domains of the chatbot. This design choice ensures tight alignment between the digital tool and the officially approved curriculum [6].

C. Ethical and Shari'ah-based Guidelines for AI in Islamic Education

The integration of AI into Islamic education cannot be separated from ethical and jurisprudential considerations. Recent discussions by Islamic juristic bodies and contemporary scholars stress the need to regulate the design and use of AI tools in light of Shari'ah objectives, scholarly traditions, and educational responsibilities [3, 7]. The International Islamic Fiqh Academy, for example, has begun to address the legal and ethical implications of AI, highlighting issues such as content authenticity, prevention of harm, respect for privacy, and the avoidance of misleading or biased outputs [7].

Within educational settings, several key principles have been proposed for AI-assisted teaching in Islamic studies: (1) ensuring the authenticity and reliability of source material, with explicit supervisory roles for qualified scholars; (2) acknowledging legitimate juristic diversity and avoiding doctrinal reductionism; (3) maintaining ethical norms in language and interaction; and (4) preserving the central role of the human teacher while positioning AI as a supportive, not substitutive, tool [3, 7]. National and institutional digital transformation strategies, such as those adopted in the United Arab Emirates, further emphasize safe, inclusive, and responsible uses of AI in higher education [8].

The present study explicitly incorporates these guidelines into the design and deployment of the chatbot. The bot's knowledge base is restricted to authenticated Hadith sources and the approved course textbook, its scope is limited to well-defined terminological and classificatory questions, and its use is embedded within a supervised university environment that includes instructor oversight and ongoing evaluation [6, 8].

By empirically examining students' experiences with this ethically constrained implementation, the study contributes to emerging frameworks for Shari'ah-compliant AI in Islamic higher education [3, 7, 8].

III. MATERIALS AND METHODS

This section describes the research context, participants, chatbot design, instruments, data collection procedures, and analytical methods used to evaluate the educational

effectiveness of the generative AI chatbot.

A. Research Design

The study adopts a convergent mixed-methods quasi-experimental design combining descriptive and quasi-experimental components. The research involved a total of 80 undergraduate students enrolled in the Hadith Sciences course, divided into two naturally occurring cohorts: a control group ($n = 40$) taught using traditional instructional methods, and an experimental group ($n = 40$) that had access to a generative AI-powered chatbot in addition to regular instruction during a six-week summer term. Quantitatively, the study compared students' performance using course grades and self-reported experiences captured through structured Likert-scale questionnaires. Qualitatively, it analyzed students' open-ended questionnaire responses, semi-structured interviews, and classroom observations to gain deeper insight into how they perceived the chatbot, integrated it into their study practices, and experienced its effects on comprehension, self-evaluation, exam preparation, and time management [3–5]. This design was selected to combine measurable outcomes with contextual qualitative evidence and to enable triangulation of students' reported learning experiences with observation notes and course performance.

This design is consistent with prior work on AI in education and Islamic pedagogy, which recommends combining quantitative outcome measures with qualitative accounts of learner experience [1–4].

B. Research Context and Participants

The research was conducted at Al Qasimia University in Sharjah, United Arab Emirates, a specialized institution focusing on Islamic studies, Arabic language, and related disciplines. The university has a diverse, international student body and a strategic interest in integrating digital technologies and smart learning solutions into its programs [8].

The Hadith Sciences course is a core requirement for students in the College of Sharia & Islamic Studies and is also taken by students from other colleges such as the College of the Holy Qur'an and College of Arts and Humanities.

The study population consisted of undergraduate students enrolled in the Hadith Sciences course during the 2024–2025 academic year. A total of 80 students were included in the analysis, divided into two naturally occurring course sections: a spring semester section that followed a traditional instructional model (control group) and a summer semester section that, in addition to traditional teaching, was given access to the generative AI chatbot over a six-week period (experimental group) [6, 8]. The sample distribution by course section is summarized in Table 1.

Table 1. Sample distribution by course section

% of Total	<i>n</i>	Instructional Model	Course Section
50	40	Traditional instruction only	Spring (Control group)
50	40	Traditional + AI chatbot (Hadith Bot)	Summer (Experimental group)
100	80	Total	

Participation in the study was voluntary, and students were informed that their questionnaire responses would be used for

research purposes with confidentiality preserved.

To minimize potential confounding effects due to different semesters, all course sections followed the same syllabus, used identical assessments and grading rubrics, and were taught by the same instructor using consistent instructional materials and pacing. Nevertheless, minor semester-specific variations in student motivation or cohort composition cannot be fully ruled out, and the study's findings are interpreted with this limitation in mind.

In addition to these two sections, an earlier autumn offering of the same course ($n = 40$) was included only in the descriptive analysis of grade distributions reported in the Results section. This autumn cohort did not have access to the chatbot and is not part of the main quasi-experimental comparison between the spring (control) and summer (experimental) sections.

C. Chatbot Design and Implementation

The educational chatbot was designed and developed by the course instructor using a generative AI platform based on large language models (ChatGPT) [2, 4]. It was configured as a course-specific conversational agent whose domain was restricted to the officially approved materials of the Hadith Sciences course. To ensure content accuracy, pedagogical suitability, and alignment with both curriculum requirements and Shari'ah standards, the chatbot content and structure underwent an expert validation process prior to implementation. The validation involved a panel of 10 university professors specialized in Prophetic Sunnah and Hadith Sciences (graduates of the College of Hadith and Islamic Studies, Islamic University of Madinah, Saudi Arabia), who reviewed the chatbot's knowledge base, terminology usage, explanations, and assessment items to ensure authenticity, doctrinal accuracy, and instructional clarity. The bot was trained and constrained using a custom knowledge base built from:

- (1) The officially adopted textbook in Hadith Sciences (*Zād al-Ṭālib min 'Ulūm al-Ḥadīth*) [6];
- (2) A curated selection of authenticated Hadith sources and classical definitions relevant to the course;
- (3) Instructor-authored explanations, examples, and formative assessment items.

The chatbot's functional structure followed four main axes reflecting the course outline [6]:

- (1) **Axis 1:** Types of Hadith by multiplicity of transmission (mutawātir; āḥād: mashhūr, 'azīz, gharīb);
- (2) **Axis 2:** Types of Hadith by end-of-chain status (qudsī, marfū', mawqūf, maqtū');
- (3) **Axis 3:** Types of Hadith by acceptance and rejection (ṣaḥīḥ, ḥasan, ḍa'īf);
- (4) **Axis 4:** Automated formative questions and practice quizzes.

In the first three axes, the chatbot provided structured explanations of key terms, clarified definitions, and offered illustrative examples, followed by brief interactive questions to reinforce understanding. In the fourth axis, it focused on summative-style practice through multiple-choice and true/false questions designed to simulate mid-term and final exam formats. The bot was integrated into the university's learning management system (Canvas) via a shared link and was also shared through course announcements, allowing

students in the experimental group to access it on demand throughout the six-week summer term [6, 8].

D. Instruments and Data Collection

Three main data sources were used:

1) Structured questionnaire

A 26-item questionnaire was developed to measure students' perceptions of the chatbot and their learning experience. It included three Likert-scale domains—Usage and Learning Behavior (5 items; five-point frequency scale: 1 = Never to 5 = Always), Understanding and Comprehension (4 items; five-point agreement scale), and Self-Evaluation (4 items; five-point agreement scale)—as well as categorical items on assessments and challenges and open-ended questions for suggestions. The questionnaire was administered to the experimental group at the end of the summer term ($n = 40$). The full instrument is provided in Appendix (Student Questionnaire). Content validity was reviewed by a panel of 10 university professors specializing in Prophetic Sunnah and Hadith Sciences (graduates of the College of Hadith and Islamic Studies, Islamic University of Madinah, Saudi Arabia), and minor wording refinements were made accordingly. Internal consistency (Cronbach's alpha) for the Likert domains was: Usage ($\alpha = 0.758$), Comprehension ($\alpha = 0.539$), and Self-Evaluation ($\alpha = 0.539$); the overall alpha across the 13 Likert items was $\alpha = 0.695$.

2) Semi-structured interviews and classroom observations

A subset of students from the experimental group ($n = 20$) participated in brief interviews during office hours to elaborate on their experiences with the chatbot. Each interview lasted approximately 15–20 min. A semi-structured interview protocol was developed prior to data collection, consisting of predefined open-ended questions focusing on students' perceptions of learning support, comprehension of Hadith terminology, usability of the chatbot, motivation, and perceived challenges. The interview guide was reviewed by subject-matter and pedagogical experts to ensure clarity and relevance, while allowing flexibility for follow-up probing questions. The instructor also recorded informal classroom observations regarding students' engagement and questions related to the tool. To enhance methodological rigor, a structured observation checklist was developed to guide the recording process, focusing on indicators such as student interaction with the chatbot, participation patterns, frequency of tool-related inquiries, and observable engagement behaviours during learning activities. Observational notes were documented systematically after each session. These qualitative data were used to contextualize and interpret the quantitative findings [3, 5].

3) Course grades and performance records

Final course grades and relevant assessment scores (e.g., mid-term and final exam results) were collected for both the control and experimental groups. While recognizing that grades are influenced by multiple factors, the study used them as an additional indicator to explore whether access to the chatbot corresponded to observable differences in academic performance [2, 4, 6]. However, given the quasi-experimental design and potential confounding variables, any improvements in grade distributions cannot be causally attributed to chatbot use. Reported grade trends are

therefore interpreted descriptively and are presented alongside students' self-reported perceptions of learning benefits, clearly distinguishing perceived learning experiences from objective performance outcomes.

E. Data Analysis

Quantitative data from the questionnaire were analyzed using descriptive statistics, including frequencies, percentages, and mean scores for each domain, in order to identify overall trends in students' perceptions and reported behaviors. Where appropriate, comparisons were made between items to highlight areas of strong perceived benefit or notable difficulty.

Course grade distributions for the control and experimental groups were compared descriptively to explore potential differences in achievement, acknowledging the limits of causal inference in quasi-experimental settings [2, 4].

Qualitative data from open-ended questionnaire responses, interviews, and classroom observations were analyzed thematically. Initial coding focused on recurring themes such as perceived advantages of the chatbot (e.g., simplification of terminology, flexibility, confidence building), challenges (e.g., question variety, technical constraints), and suggestions for improvement. These themes were then organized into higher-level categories that informed the discussion of results and the formulation of recommendations [3, 5, 6].

IV. RESULTS AND DISCUSSION

This section presents and interprets the main findings of the study regarding students' use of the chatbot, perceived learning benefits, reported challenges, and grade distributions across course sections.

Recent work on large language models in higher education suggests that tools such as ChatGPT can provide usable short-form feedback to students, while still requiring expert oversight to ensure accuracy and alignment with course objectives [12].

The results are discussed in light of prior research on AI in education and the specific pedagogical context of Hadith Sciences.

A. Patterns of Chatbot Use and Adoption

Questionnaire data from the experimental group (summer section, $n = 40$) show high levels of adoption and use of the chatbot for course-related purposes. A total of 34 students (85%) reported using the chatbot to review lecture content, clarify Hadith terminology, prepare for quizzes and the final exam, and complete course assignments.

As summarized in Table 2, 85% of students reported using the chatbot for course-related purposes, indicating high adoption and integration of the tool into their study routines. Frequent use was particularly reported in the weeks preceding major assessments, suggesting that students perceived the bot as a practical tool for structured revision.

Building on the data summarized in Table 2, Fig. 1 provides a visual representation of these quantitative results, making the patterns of chatbot use and perceived benefits easier to compare across domains.

These patterns are consistent with findings from the broader literature on AI-powered chatbots in higher education, which indicate that when such tools are closely

aligned with course requirements and easily accessible, students tend to integrate them into their regular study routines [2, 4, 5].

Table 2. Chatbot usage and perceived learning benefits (Experimental group, $n = 40$)

Domain / Indicator	n	% of students
Used the chatbot for course-related purposes	34	85
Reported improved understanding of Hadith classifications	38	95
Reported improved self-evaluation of learning	35	87.5
Used the chatbot to prepare for mid-term and final examinations	32	80
Reported better time management due to chatbot use	28	70
Recommended permanent adoption of the chatbot in the course	37	92.5
Recommended generalizing similar tools to other courses	38	95

In the context of Hadith Sciences, where dense terminology and hierarchical classifications can be challenging, the high adoption rate suggests that students recognized the chatbot as a valuable extension of classroom teaching rather than an optional add-on.

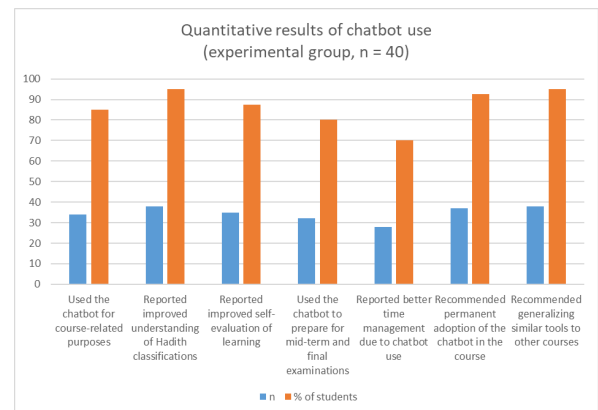


Fig. 1. Quantitative results of chatbot use among students in the experimental section ($n = 40$).

To complement the quantitative findings, qualitative data from semi-structured office-hour interviews ($n = 20$; 15–20 min each) and classroom observations provided additional context on how students experienced the chatbot in practice. Many students emphasized that the chatbot simplified dense terminology and supported conceptual differentiation within *muṣṭalaḥ al-ḥadīth*. One participant noted: “The chatbot helped me distinguish between *ṣaḥīḥ* and *ḥasan* hadith using simple explanations and examples similar to our textbook”. Another student highlighted its role in self-testing and exam preparation: “Before the final exam, I used it almost daily to test myself. When I answered incorrectly, it corrected me immediately and showed me what to revise”. Classroom observations further indicated that students increasingly referenced chatbot explanations when seeking clarification of classifications or when discussing revision strategies, suggesting that the tool was integrated into their study routines beyond individual use.

B. Effects on Comprehension, Self-evaluation, and Exam Preparation

Regarding comprehension and conceptual clarity, 38 students (95%) indicated that the chatbot helped them better

understand the main categories of Hadith, distinguish between similar terms, and link theoretical definitions to concrete examples. Students emphasized that the bot's ability to re-explain concepts in different ways contributed to simplifying abstract terminology in *muṣṭalaḥ al-ḥadīth*.

In terms of self-evaluation, 35 students (87.5%) reported that the chatbot enabled them to test their understanding, identify mistakes, and receive immediate corrective feedback. Many students indicated that interacting with the bot increased their confidence in their own level of preparedness and improved their ability to monitor their learning progress.

With respect to exam preparation and assessment, 32 students (80%) stated that the chatbot played a positive role in helping them prepare for the mid-term and final examinations. They reported that practice questions and simulated quiz items provided by the bot helped them anticipate exam formats, reinforce key concepts, and structure their study sessions more effectively. Overall, 28 students (70%) felt that the chatbot contributed to better time management and more efficient use of study hours.

Taken together, these findings suggest that the chatbot supported several dimensions of self-regulated learning—comprehension monitoring, self-testing, and strategic exam preparation—which aligns with existing research that frames AI chatbots as tools for formative assessment and personalized feedback [2, 4].

Interpreted against the study objectives, the high proportions reported for comprehension (95%), self-evaluation (87.5%), and exam preparation (80%) suggest that the chatbot functioned primarily as a formative learning scaffold—supporting repeated explanation, self-testing, and targeted revision. This pattern is consistent with prior work that frames educational chatbots as tools for self-regulated learning and metacognitive monitoring, particularly when aligned with course assessments [2, 4, 10].

In a course characterized by dense conceptual content, the ability to repeatedly revisit and practice key classifications appears to have reduced cognitive load and strengthened students' conceptual frameworks [1, 6]. From a self-regulated learning perspective, a recent systematic review concludes that educational chatbots are most often used to scaffold strategy enactment and metacognitive monitoring, while goal setting and reflection remain comparatively under-supported [10].

C. Challenges and Areas for Improvement

Despite the generally positive evaluations, some challenges were identified. Eleven students (27.5%) reported occasional difficulties in understanding certain responses from the chatbot or in dealing with technical or linguistic aspects of the system. A subset of students felt that the bot sometimes repeated similar formulations or that certain explanations were too concise and required additional elaboration. Others mentioned access limitations related to usage time or the need for a paid version to unlock extended interaction.

The reported challenges encountered by students are summarized in Table 3, which presents the frequency and percentage of students who identified each type of difficulty.

Open-ended responses highlighted several recurring suggestions for improvement, including: extending the time

limits or access quota for the chatbot; offering basic training sessions on how to use AI tools effectively; integrating audio or video explanations for complex concepts; and developing a dedicated institutional platform or app for the bot. Overall, however, students evaluated the experience very positively: 37 students (92.5%) recommended that the chatbot be adopted permanently as part of the Hadith Sciences course, and 38 students (95%) suggested that similar tools be generalized to other Islamic studies and humanities courses.

Table 3. Reported challenges in using the chatbot (Experimental group, $n = 40$)

Type of Challenge	n	% of students
Difficulty understanding some chatbot responses	11	27.5
Perceived repetition or brevity of explanations	9	22.5
Access / time-limit or quota constraints	8	20.0
Desire for more varied question types and media	10	25

Note: Some students reported more than one challenge.

To synthesize these quantitative findings, Fig. 2 presents a radar analysis of the five questionnaire domains, showing how the chatbot simultaneously supported educational engagement, understanding, self-evaluation, and exam preparation, while also revealing the main challenges reported by students.

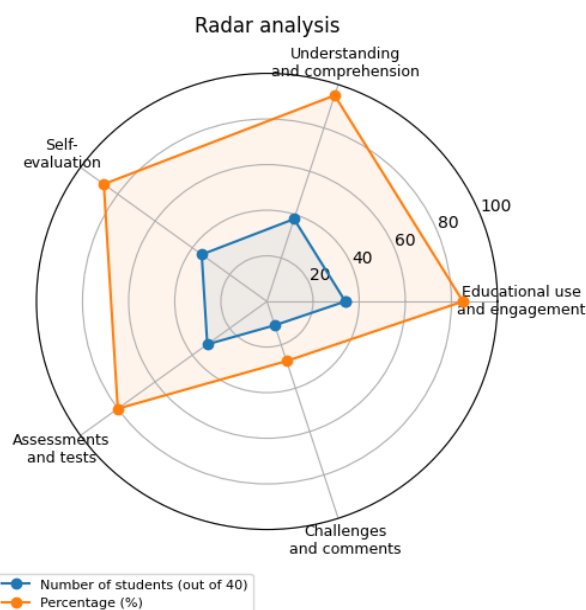


Fig. 2. Radar analysis of chatbot use and perceived learning outcomes across the five questionnaire domains (experimental group, $n = 40$).

These challenges echo concerns in the literature about the need to move beyond generic AI outputs toward tightly curated, pedagogically tuned systems [1–3]. Interview excerpts and observation records provided additional insight into these challenges, highlighting moments where students requested clarification from the instructor after receiving concise chatbot responses. Document analysis of selected interaction screenshots showed instances of repetitive phrasing, confirming students' reported experiences and supporting the interpretation of the quantitative challenge indicators.

They also underscore the importance of providing AI literacy support so that students can interpret chatbot responses critically, rather than treating them as infallible.

From a design perspective, the comments suggest that future iterations of the bot should incorporate greater variation in question types, richer explanations for difficult items, and possibly multimodal elements to support different learning styles.

D. Grade Distributions and Academic Achievement

To explore potential differences in academic achievement associated with chatbot use, grade distributions for three course sections were compared descriptively: a traditional autumn section, a traditional spring section (control group), and the experimental summer section that used the chatbot.

The analysis revealed broadly similar grade patterns between the two traditional sections, with a relatively modest proportion of students attaining the highest grade (A) and a noticeable number of students receiving low or failing grades (D or F).

In contrast, the experimental summer section showed a marked improvement in grade distribution. The proportion of students obtaining an A increased from 22.5% in the traditional sections to 42.5% in the chatbot-supported section, while the number of students receiving D or F grades decreased.

The detailed grade distributions across the three sections are summarized in Table 4.

Table 4. Grade distributions across the autumn, spring (traditional) and summer (chatbot-supported) course sections ($n = 40$ per section)

Grade	Autumn Section (Traditional)	Spring Section (Traditional)	Summer Section (chatbot)
A	8 (20%)	9 (22.5%)	17 (42.5%)
B+	8 (20%)	5 (12.5%)	6 (15%)
B	9 (22.5%)	13 (32.5%)	8 (20%)
C+	6 (15%)	7 (17.5%)	2 (5%)
C	5 (12.5%)	1 (2.5%)	7 (17.5%)
D+	1 (2.5%)	1 (2.5%)	
D	1 (2.5%)	2 (5%)	
F	2 (5%)	2 (5%)	

As shown in Table 4, the summer section that used the chatbot achieved a higher proportion of A grades and fewer low grades compared to the traditional sections. Fig. 3 illustrates the detailed grade distribution within the summer section, showing that 42.5% of students obtained an A, while the remaining grades were distributed across the B and C ranges.

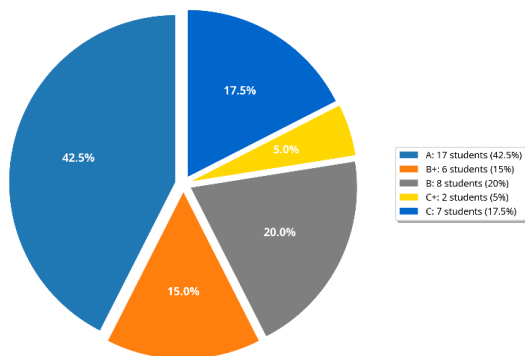


Fig. 3. Grade distribution in the chatbot-supported summer section ($n = 40$).

To provide a clearer picture of these patterns, Fig. 4 displays the grade distribution trends across the three sections (upper panel) and the percentage composition of grades within each section (lower panel).

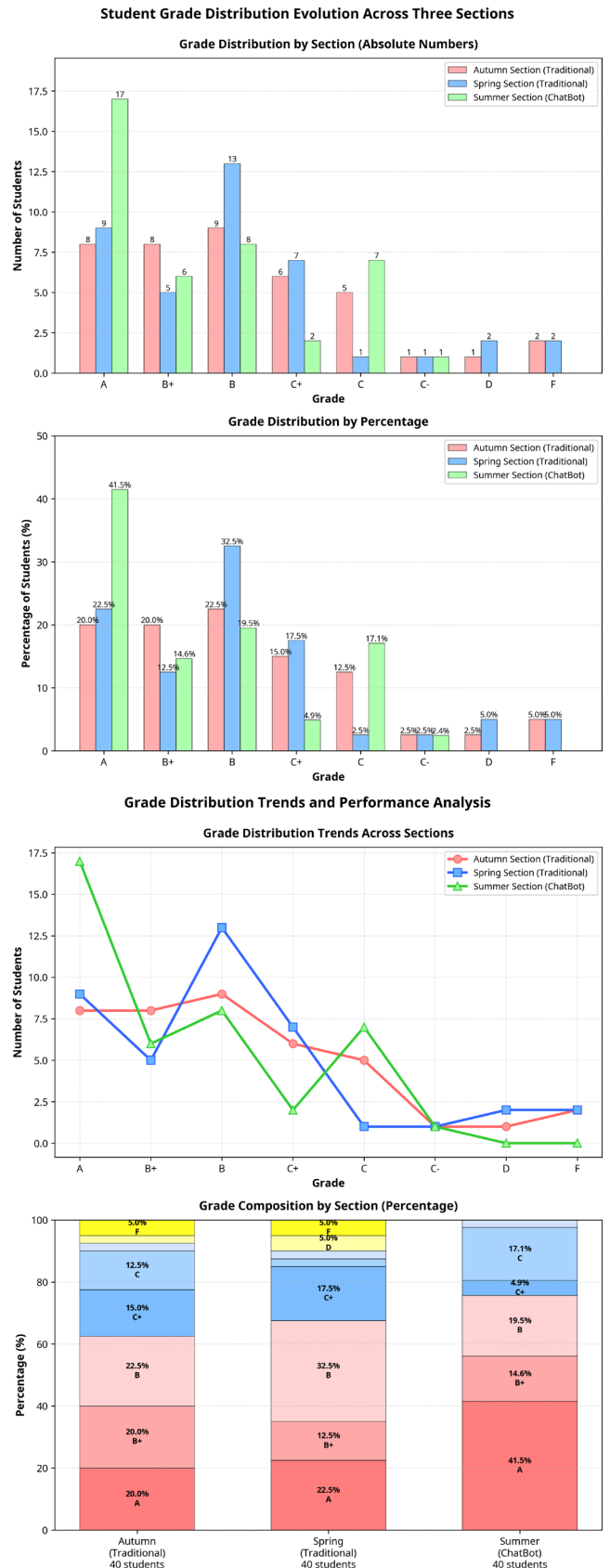


Fig. 4. Grade distributions and composition across the autumn, spring (traditional) and summer (chatbot-supported) sections.

While these results must be interpreted cautiously due to the quasi-experimental design and potential cohort differences, they align with previous findings that digital support tools can contribute to improved academic performance when they effectively scaffold understanding and exam preparation [2, 4].

In the present study, the chatbot appears to have functioned as an accessible revision companion, particularly benefiting students at the middle and lower ends of the achievement spectrum by providing extra practice and immediate clarification outside class hours.

E. Implications for AI-supported Hadith Education

Overall, the integrated findings of usage patterns, perceived learning benefits, and grade distributions suggest that generative AI-powered chatbots can play a meaningful supplementary role in teaching Hadith Sciences. The high adoption rate and strong endorsement from students indicate that the tool met real pedagogical needs that were not fully addressed by traditional instruction alone. The results also provide empirical support for the argument that AI systems, when carefully designed and ethically constrained, can be harmonized with the methodological rigor and value orientation that characterize Islamic studies [3, 6, 7].

By restricting its knowledge base to authenticated sources, aligning its structure with the official textbook, and embedding its use within an institutional framework committed to ethical AI, the present implementation responds to Shari'ah-based guidelines that call for authenticity, scholarly oversight, and preservation of the teacher's central role [3, 7, 8].

This positions the study as a concrete example of how AI can be responsibly integrated into Islamic higher education, complementing rather than displacing traditional forms of teaching and learning.

V. CONCLUSION

This study examined the effectiveness of a generative AI-powered conversational chatbot in supporting student learning in an undergraduate Hadith Sciences course at Al Qasimia University. The findings indicate that the chatbot was widely adopted by students and perceived as highly beneficial for revising course content, understanding complex Hadith classifications, engaging in self-evaluation, and preparing for examinations. Descriptive analysis of grade distributions suggests that the chatbot-supported cohort achieved higher academic outcomes than traditionally taught cohorts, although these observations are reported descriptively and are not interpreted as causal evidence.

Within the broader discourse on AI in education, the study contributes an empirically grounded model for integrating generative AI tools into Islamic studies curricula in a Shari'ah-compliant manner. It shows that AI chatbots can serve as effective supplementary tools when their design is closely aligned with the course structure, their content is rigorously supervised by specialists, and their use is embedded within clear ethical and pedagogical frameworks. The findings reported here emphasize students' perceived learning benefits—such as improved comprehension, self-evaluation, and exam preparation—while clearly distinguishing these perceptions from objective performance outcomes.

Based on the results, the following recommendations are proposed:

- (1) **Institutional Integration:** Islamic studies and humanities departments should consider integrating specialized, curriculum-aligned chatbots as supplementary tools in

courses that involve dense terminology and hierarchical classifications, such as Hadith Sciences, Fiqh, and uṣūl al-fiqh.

- (2) **Scholarly Oversight and Content Curation:** The development of educational chatbots in Islamic disciplines must remain under the supervision of qualified scholars who curate the knowledge base, validate generated content, and periodically review the system's outputs to ensure doctrinal and methodological accuracy.
- (3) **Student Training in AI Literacy:** Universities should provide basic training to help students use AI tools critically and ethically. This includes understanding the strengths and limitations of chatbots, learning how to verify AI-generated answers against trusted sources, and avoiding overreliance on automated feedback when studying or completing assignments.
- (4) **Enhancement of Assessment Features:** Future iterations of the chatbot should incorporate a wider variety of question types, adaptive difficulty levels, and richer feedback mechanisms to further strengthen its role in formative and summative assessment.
- (5) **Scaling to Other Courses and Programs:** Given the positive outcomes observed, similar chatbot models can be developed for other Islamic studies and humanities courses, contributing to a broader ecosystem of AI-enhanced learning tools tailored to Shari'ah-oriented higher education.

This study has several limitations, and future work is recommended to build on the present findings. First, the quasi-experimental design, based on naturally occurring course sections, limits the ability to draw definitive causal inferences regarding the impact of the chatbot on academic performance. Second, the sample size is modest and drawn from a single institution and course, which may restrict the generalizability of the findings. Third, the analysis of grade distributions does not fully control for potential differences in student characteristics across cohorts. Fourth, the quantitative analysis relied primarily on descriptive statistics and did not employ inferential tests (e.g., t-tests or ANOVA) to examine the statistical significance of differences between sections; therefore, observed differences should be interpreted descriptively and not as causal effects. Fifth, while interviews were conducted to capture qualitative insights, the qualitative sample was limited in scope and may not represent the full diversity of student perspectives.

Future research can address these limitations by employing more rigorous experimental or quasi-experimental designs with larger and more diverse samples, including randomized assignment where feasible. Additional studies could investigate the long-term effects of chatbot use on retention of Hadith concepts, transfer of learning to more advanced courses, and development of higher-order skills such as critical evaluation of narrations. Comparative research between AI-enhanced instruction and other digital interventions—such as learning analytics dashboards or immersive simulations—would also help clarify the distinctive contribution of generative AI chatbots to Islamic higher education. Finally, interdisciplinary collaborations between Shari'ah scholars, AI engineers, and educational researchers are needed to develop comprehensive ethical and

technical frameworks for AI in religious education.

APPENDIX: STUDENT QUESTIONNAIRE

Title of the study: Effectiveness of an AI chatbot in supporting learning in the “Hadith Sciences” course

A. Section 1: Student Information

Student name: _____

Student ID: _____

College:

- ☐ College of Sharia
- ☐ College of the Holy Qur'an
- ☐ College of Communication
- ☐ Other: _____

Year of study:

- ☐ First year
- ☐ Second year
- ☐ Third year
- ☐ Fourth year

Have you used artificial intelligence (AI) before taking this course?

☐ Yes ☐ No

How many times did you use the Hadith chatbot during this course?

- ☐ Once
- ☐ 2–4 times
- ☐ More than 5 times

Instructions: Please read each statement carefully and choose the response that best reflects your opinion or behavior.

B. Section 2: Usage and Learning Behavior

Response scale (five-point Likert): 1 = Never 2 = Rarely 3 = Sometimes 4 = Often 5 = Always

1. I used the chatbot to review the lecture content.
2. I turned to the chatbot to understand Hadith terminology (e.g., *muttaṣil*, *shādh*, *ṣaḥīḥ*...).
3. The chatbot helped me prepare for quizzes and the final exam.
4. I used the chatbot to help me complete course assignments.
5. I preferred using the chatbot instead of relying only on the textbook.

C. Section 3: Understanding and Comprehension

Response scale (five-point Likert): 1 = Strongly disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly agree

6. The chatbot helped me to simplify and better understand key concepts in Hadith terminology (*muṣṭalaḥ al-ḥadīth*).
7. The chatbot's style was easy and clear, and it explained the information in a way different from the lecture.
8. My ability to distinguish between types of Hadith (accepted, rejected, *mutawātir*, *āḥād*, etc.) improved because of using the chatbot.
9. The chatbot provided Hadith examples that were

appropriate for the theoretical concepts.

D. Section 4: Self-Evaluation

Response scale (five-point Likert): 1 = Strongly disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Strongly agree

10. By using the chatbot, I was able to evaluate my own understanding.
11. When I made mistakes in my answers, the chatbot redirected me and encouraged me when I was correct.
12. The chatbot helped me to track my weak points in a gentle and supportive way.
13. I believe that using the chatbot improved my skills in independent/self-directed study.

E. Section 5: Assessments and Tests

Response options: 1 = Don't know 2 = No 3 = Yes

14. The chatbot helped me to get ready for the mid-term exam.
15. I used the chatbot while revising for the final exam.
16. I noticed an improvement in my grades because of using the chatbot.
17. I would like the use of the chatbot to count as part of the course work/continuous assessment grade.

F. Section 6: Challenges and Comments

Response options: 1 = Sometimes 2 = No 3 = Yes

18. I found it difficult to understand some of the chatbot's responses.
19. I felt that the chatbot did not always take the *Sharī'ah* context into account in some answers.
20. Some of the information provided by the chatbot was inaccurate or needed revision.
21. I find it difficult to use the chatbot due to technical or language-related issues.

G. Section 7: Opinions and Suggestions (Open-Ended Questions)

22. What is the most important benefit you gained from using the chatbot? _____
23. What is the main weakness you noticed in the chatbot while using it? _____
24. Do you recommend adopting the chatbot on a permanent basis as part of the Hadith Sciences course? Why or why not? _____
25. Do you think this type of chatbot should be extended to other courses? Please explain.

26. Please suggest any modification or development that you think would improve the chatbot's performance in the future. _____

Thank you for your participation. All responses will be treated confidentially and used for research and quality improvement purposes only.

ETHICAL CONSIDERATIONS

The study was conducted in accordance with institutional guidelines for educational research and with *Sharī'ah*-aligned ethical principles. Students were informed about the purpose of the study, and participation in the questionnaire and interviews was voluntary. Students were assured that choosing not to participate would not affect their course evaluation or academic standing.

Confidentiality and privacy were protected by anonymizing questionnaire and interview data and reporting results in aggregate form. No personally identifying information is reported in the manuscript. Study materials and data were stored securely and accessed only by the researcher for research and quality-improvement purposes.

To support research integrity and reduce the risk of misleading outputs, the chatbot's knowledge base was restricted to authenticated, academically approved sources and remained under instructor supervision. The chatbot was used as a supplementary learning tool and did not replace the instructor's role in teaching, assessment, or mentoring.

CONFLICT OF INTEREST

The author declares no conflict of interest.

ACKNOWLEDGMENT

The author would like to thank the administration, faculty members, and students of Al Qasimia University for their support of this educational initiative and their participation in the study. Special appreciation is extended to the students of the Hadith Sciences course for their valuable feedback on the chatbot and their willingness to contribute to the advancement of AI-enhanced Islamic education. Parts of the manuscript text were drafted and language-edited with the assistance of an AI language model; all content was critically reviewed, verified, and approved by the author.

REFERENCES

- [1] B. Boulay, "Artificial intelligence in education and ethics," in *Handbook of Open, Distance and Digital Education*, O. Zawacki-Richter and I. Jung, Eds. Singapore: Springer, 2023, pp. 93–108. doi: 10.1007/978-981-19-2080-6_6
- [2] C. McGrath, A. Farazouli, and T. Cerratto-Pargman, "Generative AI chatbots in higher education: A review of an emerging research area," *Higher Education*, vol. 89, pp. 1533–1549, 2025. doi: 10.1007/s10734-024-01288-w
- [3] M. Mahmudulhassan, M. Muthoifin, and S. Begum, "Artificial intelligence in multicultural Islamic education: Opportunities, challenges, and ethical considerations," *Solo Universal Journal of Islamic Education and Multiculturalism*, vol. 2, no. 1, pp. 19–26, 2024. doi: 10.61455/sujiem.v2i01.114
- [4] M. S. Djazilan, A. Rulyansah, and J. Rihlah, "Why AI is essential for the future of Islamic education: A call for ethical and effective implementation," *EDUKASIA: Jurnal Pendidikan dan Pembelajaran*, vol. 5, no. 2, pp. 201–216, 2024. doi: 10.62775/edukasia.v5i2.1373
- [5] F. M. Othman, L. A. Majid, N. N. Amran, M. A. Nazri, and M. F. M. Ariffin, "The use of technology in Hadith pedagogy," *International Journal of Academic Research in Progressive Education and Development*, vol. 13, no. 2, pp. 94–109, 2024. doi: 10.6007/IJARPED/v13-i2/21074
- [6] S. Zakhinini, *Student's Provision in Hadith Sciences*, Sharjah, United Arab Emirates: Al Qasimia University, 2022.
- [7] International Islamic Fiqh Academy. (2025). Resolution No. 258 (3/26): Artificial Intelligence: Its Rulings, Guidelines, and Ethics. [Online]. Available: <https://iifa-aifi.org/en/56035.html>
- [8] Al Qasimia University. (2024). About Al Qasimia University and its digital transformation strategy. [Online]. Available: <https://www.alqasimia.ac.ae/>
- [9] C. Stöhr, A. W. Ou, and H. Malmström, "Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study," *Computers and Education: Artificial Intelligence*, vol. 7, 100259, 2024. doi: 10.1016/j.caeai.2024.100259
- [10] R. Guan, M. Raković, G. Chen, and D. Gašević, "How educational chatbots support self-regulated learning? A systematic review of the literature," *Education and Information Technologies*, vol. 30, pp. 4493–4518, 2025. doi: 10.1007/s10639-024-12881-y
- [11] C. Papakostas, "Artificial intelligence in religious education: Ethical, pedagogical, and theological perspectives," *Religions*, vol. 16, no. 5, 563, 2025. doi: 10.3390/rel16050563
- [1] W. Dai, J. Lin, H. Jin, T. Li, Y.-S. Tsai, D. Gašević, and G. Chen, "Can large language models provide feedback to students? A case study on ChatGPT," in *Proc. 2023 IEEE Int. Conf. Advanced Learning Technologies (ICALT)*, 2023, pp. 323–325. doi: 10.1109/ICALT58122.2023.00100

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](#)).