Informatics in Education, 2025, Vol. 24, No. 4, 737–760

DOI: 10.15388/infedu.2025.25

Real-World AI Ethics Preparation: A Debate and Case-Study Platform

Nesma BARA
University of Abderrahmane Mira-Béjaia, Faculty of Languages and Arts, Department of English, Algeria
e-mail: nesma.bara@univ-bejaia.dz

Received: February 2025

Abstract. The rapid integration of Artificial Intelligence (AI) into society demands a new generation of professionals skilled in navigating its complex ethical dimensions. This design-based research study investigates the effectiveness of 'EthicsDebateAI,' a bespoke online platform developed to address this need. Implemented through a multi-session workshop with third-year computer science undergraduates in Algeria, the intervention uses a Structured Controversy debate framework and authentic case studies to cultivate critical ethical reasoning. The intervention's impact was assessed using a mixed-methods approach. A custom pre/post-assessment revealed a significant improvement in ethical reasoning ($t(49) = 45.50$, $p < .001$, $d = 2.28$), while high student satisfaction ($M = 4.48$) and perceived career preparedness ($M = 4.44$) were confirmed via survey data. Thematic analysis of student reflections further illuminated the development of nuanced analytical skills. Findings demonstrate that targeted, interactive interventions can effectively bridge the gap between theoretical knowledge and the practical ethical competencies required of future AI professionals.

Keywords: AI ethics, structured controversy, case-based learning, ethical reasoning, career preparation, design-based research, computer science education.

1. Introduction

The pervasive integration of Artificial Intelligence (AI) is fundamentally reshaping industries and societal functions, with organizations increasingly relying on intelligent systems for strategic operations and critical decision-making (Wamba-Taguimdje *et al.*, 2020). This rapid deployment, however, is accompanied by profound ethical challenges, including algorithmic bias that perpetuates societal inequalities (O'Neil, 2016), threats to data privacy, and accountability deficits in opaque "black box" systems. While the need for ethically grounded development is clear, a significant gap persists between this demand and the practical training provided to the next generation of technology leaders.

Navigating this complex ethical landscape requires professionals equipped not just with technical expertise but with robust skills in applied ethical reasoning and critical analysis. Traditional computer science curricula, however, often treat ethics as a peripheral topic, failing to bridge the crucial gap between abstract principles and the messy, contextual dilemmas of professional practice (Hagendorff, 2020). This educational shortfall creates a risk of producing a workforce unprepared for the societal consequences of their creations, a concern particularly acute in emerging technology hubs like Algeria, where cultivating an ethically proficient workforce is critical for responsible innovation and building public trust.

To address this pedagogical challenge, this study develops and evaluates 'EthicsDebateAI,' a novel online intervention designed for computer science undergraduates. Grounded in the pedagogical framework of Structured Controversy, the platform uses structured debates and authentic case studies to move students beyond passive learning towards active, collaborative ethical deliberation. The primary goal of this research is to investigate the effectiveness of this intervention in a real-world educational setting. We aim to determine whether this targeted, dialogic approach can enhance measurable ethical reasoning skills and bolster students' perceived preparedness for the professional challenges they will inevitably face.

This study is guided by the following research questions:

RQ1 To what extent does participation in the EthicsDebateAI workshop series enhance students' AI-specific ethical reasoning skills?

RQ2 How does the intervention impact students' perceived preparedness for anticipated ethical challenges in their future AI-related careers?

Based on these questions, we formulated the following hypotheses:

H1: Students participating in the EthicsDebateAI workshop series will demonstrate a statistically significant improvement in their ethical reasoning skills from pre-test to post-test.

H2: Participants will report high levels of satisfaction with the intervention and express increased confidence in their preparedness to navigate professional ethical challenges post-intervention.

2. Literature Review

This review synthesizes research on the core ethical challenges in AI and examines pedagogical approaches for addressing them. It establishes the scholarly context for the study by first outlining the problem (the ethical landscape), then exploring existing educational solutions (pedagogy), and finally justifying the specific methods chosen (debate and case studies) through the lens of established learning theories.

2.1. *The Ethical Landscape of Artificial Intelligence*

The integration of AI systems presents a multifaceted ethical landscape. Key concerns include algorithmic bias that can perpetuate societal inequalities (O'Neil, 2016), the opacity of "black box" systems that hinders accountability (Burrell, 2016), and the societal disruption caused by AI-driven automation (Wamba-Taguimdje *et al.*, 2020). While numerous initiatives have proposed high-level ethical principles to guide development (Jobin *et al.*, 2019), a significant gap persists between these abstract guidelines and their consistent application in industry.

Mittelstadt (2019) argues that this gap is compounded by a lack of formal professional structures in AI development comparable to those in established fields like medicine or law. Furthermore, AI ethics must transcend individual conduct to address organizational pressures and institutional contexts, which heavily influence developer actions. A critical danger is "technological solutionism" the tendency to treat deep-seated societal issues as mere technical flaws amenable to an engineering fix, rather than as complex challenges requiring ongoing deliberation (Mittelstadt, 2019). Preparing AI professionals therefore requires equipping them for this nuanced, process-oriented ethical reality.

2.2. *Pedagogical Approaches to AI Ethics Education*

Addressing AI's ethical landscape requires pedagogical innovation. Traditional approaches in computer science education, often reliant on lectures and theoretical discussions, frequently fall short in preparing students to apply abstract principles to real-world dilemmas (Long and Magerko, 2020). This gap can inhibit the development of crucial professional competencies such as ethical sensitivity, reasoned judgment, and perspective-taking (Crossan *et al.*, 2013).

Effective ethics education must cultivate not only an awareness of principles but also a sense of professional responsibility and the cognitive skills needed for nuanced decision-making (Mittelstadt, 2019). This necessitates a shift towards active learning pedagogies where students must analyze trade-offs, justify choices, and engage with the technical realities of "Actually Existing AI" (Mason, 2018). Recognizing that AI systems are statistical models reliant on potentially biased data and are often brittle or prone to unforeseen failures is crucial for responsible practice (Mitchell, 2019). Thus, advancing AI ethics education demands strategies that integrate active learning with realistic, case-based reasoning.

2.3. *Structured Controversy as a Framework for Ethical Deliberation*

To bridge the gap between theory and practice, this research employs Structured Controversy, a cooperative learning framework developed by Johnson and Johnson (1979, 2009). Unlike competitive debate, which fosters a win-lose dynamic, Structured Controversy is an intellectual conflict procedure where student teams are required to research and argue

multiple sides of an issue, engage in mandatory perspective-reversal, and ultimately seek a consensus-based synthesis that integrates the best evidence from all views.

This methodology has been shown to promote higher-level reasoning, more positive peer relationships, and greater perspective-taking skills than traditional debate or individualistic learning (Johnson, Johnson, and Smith, 2000). For instance, it has been effectively used to teach complex socio-scientific issues in science education (Zohar and Nemet, 2002) and to navigate contentious topics in social studies classrooms (Hess, 2004). By adapting this established framework for AI ethics, our intervention aims to move students beyond simple argumentation towards collaborative and nuanced ethical deliberation.

2.4. *Theoretical Framework*

This pedagogical approach is grounded in two key learning theories. First, Social Constructivism (Vygotsky, 1978) posits that learning is an active, social process where individuals co-construct knowledge through interaction. The EthicsDebateAI platform embodies this theory by requiring students to build their understanding not through passive reception, but through structured dialogue, argumentation, and collaborative analysis.

Second, Communities of Practice (CoP) theory (Lave and Wenger, 1991) conceptualizes learning as situated social participation. The intervention is designed to foster a nascent CoP around AI ethics by providing a shared domain (AI's ethical challenges), a community setting for peer learning (the online platform), and a developing practice (ethical analysis and argumentation). Through these shared activities, the platform seeks to cultivate not only individual reasoning skills but also a collective identity and shared understanding related to responsible AI development.

3. Methodology

This study employs a Design-Based Research (DBR) methodology to systematically develop, implement, and evaluate the 'EthicsDebateAI' intervention within an authentic educational context (Anderson and Shattuck, 2012). DBR is particularly well-suited for this research as it provides a pragmatic framework for addressing the complex problem of enhancing AI ethics preparedness through iterative cycles of design, implementation, analysis, and refinement in a real-world setting.

3.1. *The EthicsDebateAI Intervention*

The core of this study is the EthicsDebateAI intervention, delivered to students through a 'Career-Focused Workshop Series.' The intervention operationalizes the study's theoretical frameworks by scaffolding student learning through a bespoke online platform designed to facilitate Structured Controversy and collaborative analysis.

3.1.1. *Platform Architecture and Operational Procedure*

EthicsDebateAI is a web-based application architected to guide students through a multi-stage learning process. It comprises three primary components:

- **Case-Study Repository:** A curated module housing authentic, contemporary AI ethics dilemmas (e.g., algorithmic hiring bias, privacy implications of health apps) that serve as the foundation for all activities.
- **Structured Debate Module:** The core interactive component, providing dedicated workspaces for teams to formulate arguments, engage in timed debates, and collaboratively synthesize their findings.
- **Reflection Portfolio:** An integrated tool allowing students to electronically submit and store structured written reflections, creating a longitudinal record of their intellectual and ethical development.

The operational procedure for a single learning cycle on the platform follows a clear, pedagogically driven sequence, as illustrated in Fig. 1. This student journey is designed to foster deep cognitive engagement:

1. **Case Introduction & Stance Assignment:** Students are presented with a case study and organized into small teams. The platform then assigns each team a specific stance to defend (‘pro’ or ‘con’), deliberately compelling them to engage with perspectives they might not initially hold.
2. **Argument Formulation:** Within a private team workspace, students utilize curated resources to collaboratively research their assigned stance and construct their strongest arguments.
3. **Open Debate & Rebuttal:** Teams are paired with an opposing team and engage in a shared, timed forum where they present their case, challenge opposing viewpoints, and defend their position.
4. **Perspective Reversal:** In this critical phase, the platform prompts teams to “switch sides.” They must accurately articulate the opposing team’s best arguments in their own words, a step designed to ensure deep comprehension of the alternative perspective.
5. **Synthesis & Consensus:** Finally, teams abandon their assigned roles and work collaboratively to co-author a single synthesis statement that identifies common ground, acknowledges valid points from both sides, and proposes a nuanced, consensus-based recommendation.

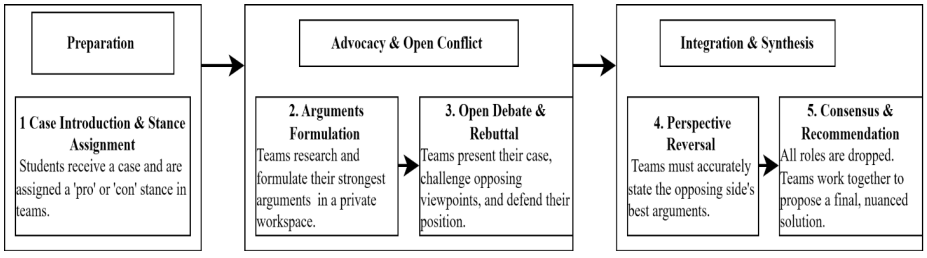

Fig. 1. Operational flowchart of the student learning cycle within the EthicsDebateAI platform, mapping the process to the key phases of Structured Controversy.

3.1.2. *Pedagogical Implementation of Structured Controversy*

The platform's operational procedure is a direct digital implementation of the Structured Controversy framework (Johnson and Johnson, 2009). Unlike traditional debate, which can foster a zero-sum, win-lose dynamic, this cooperative learning model is explicitly designed to ensure students master multiple perspectives before seeking a synthesized resolution. The core phases of the model are embedded into the platform's workflow: (1) learning and presenting a position (our Argument Formulation and Open Debate phases); (2) arguing from the opposing perspective (our Perspective Reversal phase); and (3) reaching a consensus through synthesis (our Synthesis & Consensus phase). By digitally scaffolding this process, the EthicsDebateAI platform guides students from simple advocacy to complex, collaborative ethical reasoning.

3.2. *Research Design and Participants*

The study employed a single-group pretest-posttest design embedded within the broader DBR framework. The DBR process involved distinct, interconnected stages: (1) Problem Analysis to define AI ethics education needs at the university; (2) Intervention Design of the EthicsDebateAI platform and workshop activities; (3) Implementation with the target student population; (4) Evaluation and Analysis of learning outcomes; and (5) Refinement based on empirical findings for future iterations.

Participants were third-year undergraduate computer science students (N = 50) at Ferhat Abbes University, Setif 1, Algeria. This cohort was selected for their foundational AI knowledge and proximity to career entry, making the intervention maximally relevant. Recruitment for the voluntary 'Career-Focused Workshop Series' utilized convenience sampling via course announcements. The principal researcher's background in English for Specific Purposes (ESP) informed the pedagogical emphasis on argumentation and communication, complementing the students' technical training. Participation was voluntary, and informed consent was obtained from all students prior to the start of the study. All study materials and interactions were conducted in English.

3.3. *Data Collection and Instruments*

A mixed-methods approach was used to gather comprehensive data throughout the 'Career-Focused Workshop Series,' which consisted of six 90-minute sessions held weekly over six weeks. The following instruments were employed:

- **AI Ethical Reasoning Assessment (Tool 1; Appendix A):** A custom pre/post-test instrument was administered electronically to measure changes in ethical reasoning. Inspired by established measures (e.g., the Defining Issues Test; Rest *et al.*, 1999), it presented realistic AI dilemmas requiring students to decide and rate the importance of various ethical considerations. To ensure content validity,

the instrument was reviewed by a panel of three experts in AI ethics and computer science education. It was then pilot-tested with a separate group of students (n = 15) to refine item clarity. For the present study, the scale demonstrated strong internal consistency (Cronbach's $\alpha = .86$).

- **Structured Reflections (Tool 2; Appendix B):** At key junctures, students submitted structured written reflections via the platform's Reflection Portfolio. Prompts were designed to elicit their evolving understanding of AI ethics, analytical challenges, and changes in perspective.
- **Post-Workshop Satisfaction Survey (Tool 3; Appendix C):** An anonymous online survey administered at the workshop's conclusion used a combination of 5-point Likert-scale and open-ended questions to assess overall satisfaction, platform usability, perceived learning impact, and career preparedness.
- **Platform Engagement Data:** The platform automatically logged quantitative metrics of student activity (e.g., module completion rates, time on platform, number of discussion contributions) to provide objective measures of engagement.

3.4. *Data Analysis*

Data were analyzed systematically using a convergent mixed-methods design, allowing for the integration of quantitative and qualitative findings. All quantitative analyses were performed using IBM SPSS Statistics (Version 26).

- **Quantitative Analysis:** To address RQ1, pre- and post-intervention scores from the AI Ethical Reasoning Assessment (Tool 1) were compared using a paired-samples t-test, with effect size calculated using Cohen's d. Descriptive statistics (means, standard deviations) were used to summarize survey data (Tool 3) and platform engagement metrics. To explore the relationship between engagement and learning, Pearson correlation coefficients were computed between platform usage data and the gains in assessment scores.
- **Qualitative Analysis:** All qualitative data from structured reflections (Tool 2) and open-ended survey questions (Tool 3) were analyzed concurrently using Thematic Analysis (Braun and Clarke, 2006). This involved a rigorous six-phase process: (1) data familiarization, (2) systematic initial coding, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. To ensure trustworthiness, two researchers independently coded a 20% subset of the data, achieving high inter-rater reliability (Cohen's Kappa = 0.86). Discrepancies were resolved through discussion to refine the final coding framework.
- **Data Integration:** Finally, findings from the quantitative and qualitative analyses were triangulated during the interpretation phase to develop a comprehensive and robust understanding of the EthicsDebateAI intervention's impact.

4. Results

This section presents the findings from quantitative and qualitative data analyses, structured to address the study's primary research hypotheses.

4.1. *Impact on AI Ethical Reasoning Skills (Hypothesis 1)*

Hypothesis 1 predicted that students participating in the EthicsDebateAI workshop would demonstrate a significant improvement in their AI-related ethical reasoning skills. To test this, scores from the AI Ethical Reasoning Assessment administered before (Pre-Test) and after (Post-Test) the intervention were compared using a paired-samples t-test for all 50 participants.

The descriptive statistics, presented in Table 1, show a substantial increase in students' ethical reasoning scores. The cohort's mean score rose from 43.98 (SD = 8.82) at pre-test to 64.12 (SD = 8.63) at post-test.

The results of the paired-samples t-test (Table 2) confirmed that this observed increase was statistically significant. The analysis revealed a mean increase of 20.14 p oints, $t(49) = 45.50$, $p < .001$. The 95% confidence interval for this difference [19.25, 21.03] does not contain zero, further underscoring the certainty of this result. The effect size for this improvement was exceptionally large (Cohen's $d = 2.28$), indicating that the intervention had a profound and practically significant impact on participants' ethical reasoning capabilities.

These results provide robust quantitative support for Hypothesis 1, demonstrating that the EthicsDebateAI workshop was highly effective in enhancing students' AI ethical reasoning skills.

Table 1

Descriptive Statistics for AI Ethical Reasoning Scores (Pre- and Post-Intervention, N = 50)

	Mean	N	Std. Deviation	Std. Error Mean
Pre	43.98	50	8.817	1.247
Post	64.12	50	8.628	1.220

Table 2

Paired Samples T-Test Comparing Pre- and Post-Intervention AI Ethical Reasoning Scores (N = 50)

Pair 1	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
AI_Reasoning_Pre_Score – AI_Reasoning_Post_Score	-20.140	3.130	.443	-21.030	-19.250	-45.500	49	.000

**Note. Cohen's d calculated as Mean Difference / SD_Pre_Score.*

4.2. *Student Satisfaction and Perceived Career Preparedness (Hypothesis 2)*

Hypothesis 2 proposed that students report high levels of satisfaction and feel more prepared for future ethical challenges. This was assessed using the Post-Workshop Satisfaction Survey. Likert-scale items were grouped into four conceptual subscales, and a mean score was computed for each participant on each subscale.

The findings, summarized in Table 3, show consistently high levels of agreement across all dimensions. To address the reviewer's query regarding non-integer minimum values, a note has been added to clarify that these values represent the lowest observed mean subscale score from any single participant, not the lowest point on the 1–5 scale itself.

The high mean scores, all well above the "Agree" threshold of 4.0, indicate:

- **Strong Usability and Engagement (M = 4.28):** Students found the EthicsDebateAI platform easy to use and the activities compelling.
- **Significant Perceived Learning (M = 4.42):** Participants strongly agreed that the workshop enhanced their understanding of AI ethics and their analytical abilities.
- **High Career Relevance (M = 4.44):** Students perceived the content as highly relevant and effective in preparing them for ethical challenges in the workplace.
- **Excellent Overall Satisfaction (M = 4.48):** Participants were overwhelmingly positive about the experience and would recommend it to peers.

These quantitative findings provide strong support for Hypothesis 2.

4.3. *Qualitative Feedback on Student Experience*

To add depth and context to the high satisfaction ratings, open-ended survey responses were subjected to thematic analysis. This process yielded four predominant themes that illuminate the student experience and the pedagogical mechanisms driving the intervention's success, as summarized in Table 4.

Table 3

Mean Scores and Standard Deviations for Satisfaction Survey Subscales (N = 50)

Subscale	Number of Items	Mean	Std. Deviation	Minimum	Maximum
A: Platform Usability & Engagement	5	4.28	0.55	3.20	5.00
B: Learning & Skill Development	5	4.42	0.52	3.60	5.00
C: Relevance & Career Preparedness	5	4.44	0.53	3.50	5.00
D: Overall Satisfaction	2	4.48	0.57	3.50	5.00

Note. Subscale scores are the mean of items (rated 1=Strongly Disagree to 5=Strongly Agree) for each participant. The 'Minimum' and 'Maximum' columns reflect the lowest and highest observed mean subscale scores from any single participant, clarifying why non-integer values are present.

Table 4

Thematic Schema and Illustrative Quotes from Open-Ended Survey Responses (N=50)

Main Theme	Definition / Focus	Example Illustrative Quote
1. **Valued Pedagogical Approaches**	Students identified specific workshop elements as particularly helpful for learning.	"*The structured controversy debates really forced me to see both sides of an issue... much better than just being told what to think.*" (P17) "*The case studies were engaging because they felt like real problems we might actually face*." (P9)
2. **Perceived Learning and Skill Gains**	Participants articulated specific improvements in their skills, knowledge, or awareness.	"*My ability to formulate a reasoned ethical argument definitely improved. Before, I wasn't sure how to structure my thoughts.*" (P24) "*I feel much better equipped now to spot potential bias in algorithms.*" (P11)
3. **Enhanced Career Relevance and Impact**	The workshop influenced students' perspectives on their future professional roles.	"*This workshop made me realize that ethics isn't just an afterthought; it's a fundamental part of being a responsible AI professional.*" (P3) "*I feel more confident now that I could speak up if something felt ethically wrong.*" (P45)
4. **Constructive Suggestions for Improvement**	Participants provided specific feedback for enhancing the workshop or platform.	"*Maybe some case studies could include more technical details...so we can see how the ethical issue links to the code.*" (P19) "*It would be good to have more time for the synthesis part of the debate.*" (P36)

The qualitative data reveal a clear consensus among participants regarding the value of the intervention's active learning design. As captured by P17 and P9, students specifically credited the Structured Controversy debates and authentic case studies as the primary drivers of their learning. The themes of «Perceived Learning and Skill Gains» and «Enhanced Career Relevance» are particularly telling. Participants articulated a tangible progression in their abilities, moving from uncertainty to confidence in formulating and structuring ethical arguments (P24).

Crucially, the feedback underscores the intervention's success in forging a strong link between abstract ethical theory and concrete professional practice. The sentiment expressed by P3 that ethics is «a fundamental part of being a responsible AI professional» signals a profound shift in perspective. This, coupled with the sense of newfound professional agency articulated by P45 («I feel more confident now that I could speak up»), provides strong qualitative evidence that the workshop not only developed skills but also began to cultivate a professional ethical identity.

4.4. *Developmental Trajectory: Insights from Student Reflections*

The thematic analysis of the four structured reflections submitted sequentially throughout the workshop chronicled a clear developmental journey in students' ethical reasoning. The predominant theme identified was a cognitive shift from **Surface-Level Awareness to Deep, Nuanced Understanding**.

Early reflections often framed ethical issues as simple binary choices with clear right or wrong answers. In contrast, later reflections demonstrated a sophisticated appreciation for complexity, trade-offs, and the contextual nature of ethical dilemmas.

Students increasingly acknowledged the competing values at stake and the difficulty of finding a perfect solution. One student's narrative captured this evolution with remarkable clarity:

> "*Before the debate on facial recognition, I mostly thought about its convenience for things like unlocking my phone. Now, after having to argue against it, I'm much more aware of the potential for surveillance, misidentification, and how it could be used in ways that harm certain communities.*" (P21)

This quote exemplifies a common pattern: the structured debate process, particularly the perspective-reversal phase, compelled students to move beyond their initial assumptions. These narrative accounts offer compelling, human-centered evidence of the cognitive growth fostered by the workshop, providing a rich context that validates and explains the significant quantitative improvements reported previously.

4.5. *Platform Usage and Engagement Patterns*

Platform usage data provided objective measures of student engagement, indicating high and consistent participation across the cohort. On average, students completed 85.3% of the core modules and spent nearly eight hours actively engaged with the platform's content and debate forums.

To investigate the relationship between the volume of student activity and learning outcomes, correlational analyses were conducted. These analyses explored the intercorrelations among engagement metrics and their relationship with post-intervention ethical reasoning scores. Pearson correlation coefficients (r) are summarized in Table 6.

A critical finding emerged from this analysis: there was no statistically significant linear relationship between the *quantity* of student engagement (e.g., time on platform, number of posts) and their final ethical reasoning scores. While all engagement metrics were, as expected, strongly inter-correlated ($r > .93$, $p < .01$), the sheer volume of a student's activity did not predict the magnitude of their learning gain.

This counterintuitive result, particularly in the context of the large overall learning gains, carries significant pedagogical implications. It suggests that the quality of en-

Table 5

Selected Platform Engagement Metrics (N = 50)

Descriptive Statistics					
Metrics	N	Minimum	Maximum	Mean	Std. Deviation
Modules Completed (%)	50	60	100	85.28	11.082
Time_Platform_Hours	50	4.5	12.3	7.868	1.6211
Forum_Posts	50	1	10	5.46	2.435

Note: Data derived from platform logs.

Table 6

Pearson Correlation Between Percentage of Modules Completed and Post-Intervention AI Ethical Reasoning Score (N = 50)

Variable	1. Modules Completed (%)	2. Forum Posts	3. Time on Platform (Hrs)	4. AI Reasoning Post-Score
1. Modules Completed (%)	---	.972**	.934**	-.227
2. Forum Posts		---	.947**	-.183
3. Time on Platform (Hrs)			---	-.218
4. AI Reasoning Post-Score				---

*** Correlation is significant at the 0.01 level (2-tailed).*

gagement prompted by the structured activities, rather than the mere volume of interaction, was the critical driver of learning. The intervention's success was not a function of "more is better." Instead, a general threshold of active participation appeared sufficient for students to benefit, with the pedagogical architecture of the Structured Controversy framework being the primary mechanism for improvement. This finding points to the primacy of instructional design over raw interaction time in fostering complex cognitive skills.

5. Discussion

This study designed, implemented, and evaluated 'EthicsDebateAI,' a bespoke online platform aimed at enhancing the AI ethics preparedness of undergraduate computer science students. Grounded in a DBR methodology and centered on the pedagogical framework of Structured Controversy, the intervention was delivered via a 'Career-Focused Workshop Series' at Ferhat Abbes University, Setif 1. The findings not only confirm the intervention's efficacy but also offer critical insights into *how* active, dialogic learning can bridge the well-documented gap between abstract ethical principles and applied professional reasoning.

5.1. *Answering Research Questions and Confirming Hypotheses*

The investigation's results provide compelling, multi-faceted answers to the guiding research questions and robustly support both of the study's hypotheses.

RQ1 and H1: Enhancing AI-Specific Ethical Reasoning

In response to RQ1, which queried the intervention's impact on ethical reasoning, the findings provide a clear and resounding affirmative, robustly confirming Hypothesis 1. The exceptionally large effect size (d = 2.28) offers quantitative proof of a profound shift in students' analytical capabilities. However, it is the qualitative data that reveal

the pedagogical mechanism driving this change, offering an empirical counterpoint to the critiques of traditional, passive AI ethics education (e.g., Long and Magerko, 2020), which often fails to move students beyond principle recognition to applied judgment.

The key mechanism appears to be the Structured Controversy framework itself. As highlighted by P17 ("forced me to see both sides"), the mandatory perspective-reversal phase is particularly critical. This step directly compels students to dismantle their own biases, engage with the strongest arguments of an opposing view, and collaboratively seek a nuanced synthesis. This process actively fosters cognitive flexibility and sophisticated perspective-taking skills that Crossan *et al.* (2013) identify as crucial for mature ethical judgment. By designing an experience where students must actively construct and deconstruct arguments, the intervention successfully cultivated the very skills that passive learning struggles to impart.

RQ2 and H2: Perceived Preparedness for Professional Ethical Challenges

Similarly, the study provides a strong positive answer to RQ2, comprehensively supporting Hypothesis 2. The high mean survey scores for 'Relevance & Career Preparedness' (M = 4.44) provide quantitative confirmation.

More importantly, the thematic analysis suggests this readiness transcends mere academic confidence; it reflects a budding sense of professional identity and moral agency, a crucial element that Mittelstadt (2019) notes is often underdeveloped in technical training. When a student states, "I feel more confident now that I could speak up if something felt ethically wrong" (P45), it signals a shift from being a passive recipient of ethical codes to an active agent prepared to enact them. This newfound agency is a powerful antidote to the "technological solutionism" that Mittelstadt (2_019) warns against. The intervention helped students see ethics not as a technical problem to be fixed with a better algorithm, but as a complex human and societal challenge requiring ongoing deliberation and professional courage.

5.2. *A Critical Insight: The Primacy of Pedagogical Design*

Beyond confirming the hypotheses, a particularly noteworthy finding is the lack of a linear correlation between the quantity of platform engagement and the magnitude of learning gains. This suggests the intervention's success was not a function of "more is better" but rather a testament to the primacy of pedagogical design. The structured, goal-oriented nature of the debate and reflection activities ensured that any level of sincere participation was meaningful.

This result challenges a common assumption in educational technology that equates more interaction time with deeper learning. For complex, higher-order domains like ethics, this study suggests that the architectural quality of the learning experience is structured to provoke critical thinking, perspective-reversal, and synthesis is a far more potent driver of learning than the sheer volume of student activity.

5.3. *Implications of the Study*

The findings carry significant implications for AI ethics education, educational technology design, and computer science curriculum development:

- **For Pedagogy:** This study provides strong empirical evidence for the efficacy of active, dialogic learning methodologies over passive instruction for teaching AI ethics. It specifically validates the adaptation of Structured Controversy to an online, asynchronous platform as a highly effective model for cultivating nuanced ethical reasoning.
- **For Curriculum Development:** The high perceived career relevance underscores the urgent need to embed practical, interactive ethics training as a core component within technical curricula, rather than an optional afterthought. The EthicsDebateAI model serves as a validated, adaptable blueprint for universities seeking to better prepare their students for the demands of the modern technology industry.
- **For Global Education:** By developing and validating this intervention in an Algerian context, this research demonstrates the transportability and importance of such educational models beyond traditional Western-centric institutions. It highlights the need to design interventions that are culturally and professionally resonant with diverse local student populations.
- **For Educational Technology:** The DBR approach was instrumental in creating a tool that is both theoretically sound and deeply responsive to learner needs. Student feedback, such as the desire for more technical details in case studies, provides direct, actionable guidance for the next design cycle, reinforcing the value of DBR in educational innovation.

5.4. *Limitations and Future Directions*

While the study yields valuable insights, its limitations must be acknowledged to contextualize the findings and guide future research.

First, the single-group pretest-posttest design, while appropriate for this initial DBR cycle, cannot definitively establish causality. Although the large effect size is compelling, the design cannot entirely rule out the influence of confounding variables, such as students' concurrent coursework or general maturation. Therefore, the next logical research step is to employ a quasi-experimental design with a control group to isolate the intervention's effects more rigorously.

Second, the findings' generalizability is constrained by the specific sample of third-year computer science students from a single Algerian university. The use of convenience sampling may also introduce a self-selection bias, as students who volunteered may have been more motivated than their peers. Future studies should seek to replicate these findings with more diverse student populations across different institutions and cultural contexts to establish the broader applicability of the model.

Finally, while the custom-developed AI Ethical Reasoning Assessment was rigorously designed and demonstrated strong internal consistency, it would benefit from further

psychometric validation with larger, more diverse samples. As an output of this DBR study, the findings reflect the intervention at its current stage; ongoing refinement based on these results is an inherent and planned part of the research process.

6. Conclusion and Future Directions

This study designed, implemented, and empirically validated 'EthicsDebateAI,' establishing its efficacy as a pedagogical tool for enhancing AI-specific ethical reasoning and career preparedness among computer science undergraduates. The findings provide a robust endorsement for active, context-rich learning methodologies. We have demonstrated that a digitally-scaffolded Structured Controversy framework, grounded in authentic case studies, is a powerful instrument for equipping future technology professionals to navigate the complex ethical landscape of artificial intelligence. By cultivating the capacity for applied ethical reasoning, interventions such as EthicsDebateAI represent a vital contribution to the global effort to foster a more responsible and socially beneficial technological future.

Building upon this study's foundation, several promising avenues for future research emerge:

First, future work should prioritize methodological enhancement. The next logical step is to employ a quasi-experimental design with a control group to establish a more definitive causal link between the intervention and learning outcomes. Furthermore, longitudinal studies are needed to track the long-term impact of this training on participants' professional ethical decision-making once they enter the workforce.

Second, pedagogical refinement offers a rich area for investigation. In-depth research could delineate which specific elements of the intervention such as the perspective-reversal phase, specific case characteristics, or the nature of peer interaction most significantly drive the development of ethical reasoning. Such findings could inform the iterative evolution of the platform with new features, such as AI-driven feedback or more complex ethical simulations.

Finally, scaling and validating the intervention is a critical long-term goal. This includes adapting and testing the EthicsDebateAI model across diverse cultural, institutional, and disciplinary settings. Concurrently, the continued psychometric validation of the assessment tool with larger samples will be essential to ensure its robustness and broad applicability. Pursuing these directions will not only strengthen the claims made in this study but also maximize the potential impact of this educational model.

References

Anderson, T., Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher,* 41(1), 16–25. https://doi.org/10.3102/0013189X11428813

Anderson, T., Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher,* 41(1), 16–25. https://doi.org/10.3102/0013189X11428813

Braun, V., Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3(2),

77–101. https://doi.org/10.1191/1478088706qp063oa
Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society,* 3(1). https://doi.org/10.1177/2053951715622512
Crossan, M., Mazutis, D., Seijts, G. (2013). In search of virtue: The role of virtues, values and character strengths in ethical decision making. *Journal of Business Ethics,* 113(4), 567–581. https://doi.org/10.1007/s10551-012-1326-y
Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines,* 30(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8
Hess, D.E. (2004). Controversies about controversial issues in democratic education. *PS: Political Science & Politics,* 37(2), 257–261. https://doi.org/10.1017/S1049096504004196
Jobin, A., Ienca, M., Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence,* 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2
Johnson, D.W., Johnson, R.T. (1979). Conflict in the classroom: Controversy and learning. *Review of Educational Research,* 49(1), 51–69. https://doi.org/10.3102/00346543049001051
Johnson, D.W., Johnson, R.T. (2009). Energizing learning: The instructional power of conflict. *Educational Researcher,* 38(1), 37–51. https://doi.org/10.3102/0013189X08330540
Johnson, D.W., Johnson, R.T., Smith, K.A. (2000). Constructive controversy: The power of intellectual conflict. *Change: The Magazine of Higher Learning,* 32(1), 28–37. https://doi.org/10.1080/00091380009602704
Lave, J., Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
Long, D., Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3313831.3376727
Mason, H. (2018). *AI is not magic*. O'Reilly Radar. Retrieved from https://www.oreilly.com/radar/ai-is-not-magic/
Mitchell, M. (2019). *Artificial Intelligence: A Guide for Thinking Humans*. Farrar, Straus and Giroux.
Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence,* 1(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4
O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Books.
Rest, J., Narvaez, D., Bebeau, M.J., Thoma, S.J. (1999). *Postconventional Moral Thinking: A Neo-Kohlbergian Approach*. Lawrence Erlbaum Associates.
Vygotsky, L.S. (1978). *Mind in Society: The Development of Higher Psychological Processes* (M. Cole, V. John-Steiner, S. Scribner, E. Souberman, Eds.). Harvard University Press.
Wamba-Taguimdje, S.-L., Fosso Wamba, S., Kala Kamdjoug, J.R., Tchatchouang Wanko, C.E. (2020). Influence of artificial intelligence (AI) on firm performance: The business value of AI-based transformation projects. *Business Process Management Journal,* 26(7), 1893–1924. https://doi.org/10.1108/BPMJ-10-2019-0411
Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press.
Zohar, A., Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching,* 39(1), 35–62. https://doi.org/10.1002/tea.10008

N. Bara, Dr., is an Associate Professor of English at the University of Béjaïa, Algeria. Her research focuses on "decolonial methodologies" in critical intercultural studies, examining the intersection of language, power, and identity. Her expertise includes sociolinguistics and the transformative impact of Artificial Intelligence (AI) in educational and research settings. An active contributor to the global academic community, she is an Editorial Board Member for the Journal of Global Education and serves as a reviewer for prestigious journals including the International Journal of Applied Linguistics (Wiley), Interchange (Springer), and Frontiers.

Appendix A
AI Ethical Reasoning Assessment (Tool 1)

Purpose: To measure changes in students' ethical reasoning specifically concerning AI dilemmas before and after the intervention. This instrument is administered as a pre-test and a post-test.

Structure: The assessment presents 3–4 short AI ethical dilemmas. For each dilemma, students are asked to:

a) Indicate their judgment/decision regarding the scenario.

b) Rate the importance of several provided statements (representing different ethical considerations/reasoning levels) in making that decision.

c) Rank the top 3–4 most important statements they considered.

(The following is an example of one dilemma. The full instrument would include 2–3 additional dilemmas covering different AI ethics areas such as data privacy, autonomous systems, or AI in surveillance, each with its own set of tailored importance statements and ranking prompts.)

Dilemma Example 1: Algorithmic Hiring Tool

Scenario: A software company is developing an AI tool to screen job applications (CVs) for software engineering positions to handle a high volume of applicants efficiently. Early testing reveals the AI seems to slightly favor candidates from certain universities known historically for having predominantly male graduates, even though gender is not an explicit input variable. The project manager argues the efficiency gains are significant and the bias is minor and can perhaps be tweaked later. The lead developer is concerned about launching a tool that might perpetuate existing biases and potentially discriminate against qualified female candidates or those from less represented universities.

Decision Prompt: As the lead developer, what would be your primary recommendation regarding the launch of this AI screening tool? (Please select one):

___ Recommend launching now, focusing on efficiency, and addressing bias later.

___ I just realised there is an unnecessary newline on page 16 (see attached)cide.

Importance Rating Prompt: How important were the following considerations in making your recommendation?

(Scale: 1 = No Importance, 2 = Little Importance, 3 = Some Importance,
4 = Much Importance, 5 = Great Importance)

1. Ensuring the company meets its hiring targets quickly and efficiently.
(__ 1 __ 2 __ 3 __ 4 __ 5)

2. Avoiding potential legal challenges related to discriminatory hiring practices. (__ 1 __ 2 __ 3 __ 4 __ 5)

3. The potential negative impact on the company's public reputation if the bias becomes known. (__ 1 __ 2 __ 3 __ 4 __ 5)

4. The fundamental ethical principle of fairness and equal opportunity for all qualified applicants. (__ 1 __ 2 __ 3 __ 4 __ 5)

5. The technical challenge and feasibility of actually removing the bias from the algorithm. (__ 1 __ 2 __ 3 __ 4 __ 5)

6. The lead developer's personal responsibility to prevent harm caused by the technology they build. (__ 1 __ 2 __ 3 __ 4 __ 5)

7. Adhering strictly to the project manager's directives and company goals. (__ 1 __ 2 __ 3 __ 4 __ 5)

8. The long-term consequences for diversity within the company's workforce. (__ 1 __ 2 __ 3 __ 4 __ 5)

9. Whether similar tools used by competitors also have biases. (__ 1 __ 2 __ 3 __ 4 __ 5)

10. Upholding professional codes of conduct for software engineers regarding societal impact. (__ 1 __ 2 __ 3 __ 4 __ 5)

11. The potential benefits the tool could bring if the bias issue were resolved. (__ 1 __ 2 __ 3 __ 4 __ 5)

12. How easily external auditors or users could detect bias. (__ 1 __ 2 __ 3 __ 4 __ 5)

Ranking Prompt: From the list above, select the FOUR most important considerations in making your recommendation. Please list them below by statement number.

1. Most Important: _____

2. Second Most Important: _____

3. Third Most Important: _____

4. Fourth Most Important: _____

Appendix B
Student Reflection Prompts (Tool 2)

Purpose: To elicit qualitative data on students' learning processes, ethical sense-making, challenges encountered, and perceived changes in perspective regarding AI ethics throughout the workshop intervention.

Format: Open-ended questions presented at specific junctures within the workshop via the EthicsDebateAI platform.

A. Prompt Following Completion of a Debate Module (e.g., Facial Recognition in Public Spaces)

Instructions: Please reflect on the debate module about facial recognition you just completed. Answer the following questions thoughtfully (approx. 150–250 words total):

1. Before the debate, what was your initial stance on using facial recognition technology by authorities in public spaces?

2. Having researched and argued for both supporting and opposing viewpoints (as required by the Structured Controversy format), what is one argument from the opposing side that you now find particularly compelling or challenging? Explain why.

3. How did the process of debating this topic influence your understanding of the trade-offs between societal security/convenience and individual privacy/potential for misuse in the context of AI?

B. Prompt Following Analysis of a Case Study (e.g., Data Privacy Breach in a Health App)

Instructions: Consider the case study involving the health app data breach. Please respond to the following (approx. 150–250 words total):

1. Identify the two primary ethical principles that you believe were in conflict in this scenario. Explain the tension between them.

2. If you were the lead software engineer on this project before the breach occurred, what specific technical or procedural safeguards related to data handling would you have advocated for, and why?

3. Beyond technical solutions, what responsibility do you think the company (and its developers) had towards users regarding their data, even if legally compliant at the time?

C. Mid-Workshop Reflection Prompt (After approx. half the sessions)

Instructions: Reflecting on the workshop activities so far (debates, case studies, discussions):

1. Describe one specific AI ethics concept (e.g., transparency, accountability, fairness) that has become clearer or more complex for you. What caused this change in understanding?

2. What has been the most challenging aspect of analyzing these AI ethical dilemmas or participating in the debates? Why?

3. How, if at all, are you starting to connect these ethical discussions to the technical aspects of computer science you have learned in other courses?

D. Final Workshop Reflection Prompt (End of the Series)

Instructions: Looking back at the entire workshop series on AI ethics:

1. Describe the most significant change, if any, in how you approach thinking about the ethical implications of developing or deploying AI systems. What prompted this change?

2. Consider the ethical challenges discussed (bias, privacy, accountability, job displacement, etc.). Which one do you anticipate being most relevant or difficult to navigate in your future career as a computer science professional in Algeria or elsewhere? Explain your reasoning.

3. How has participating in this workshop, particularly using the EthicsDebateAI platform, influenced your feeling of preparedness to address ethical issues in your future work? Be specific about any skills or awareness you feel you have developed.

Appendix C
Post-Workshop Satisfaction Survey (Tool 3)

Purpose: To assess student satisfaction with the "Career-Focused Workshop Series" featuring the EthicsDebateAI platform, evaluate its perceived effectiveness in enhancing ethical understanding and skills, and gauge its impact on perceived preparedness for professional ethical challenges.

Format: Anonymous online survey administered after the final workshop session.

Instructions: Please take a few minutes to provide your honest feedback on the AI Ethics workshop series and the EthicsDebateAI platform. Your responses are anonymous and will be used to evaluate and improve this educational experience.

Part 1: Likert-Scale Questions

Instructions: Please indicate your level of agreement with the following statements using the scale provided: (Scale: 1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree)

A. Platform Usability & Engagement:

1. The EthicsDebateAI platform was easy to navigate and use.
   (__ 1 __ 2 __ 3 __ 4 __ 5)

2. The instructions for the debate modules and case studies on the platform were clear.
   (__ 1 __ 2 __ 3 __ 4 __ 5)

3. The structured debate format (Structured Controversy) was engaging.
   (__ 1 __ 2 __ 3 __ 4 __ 5)

4. The case studies presented were interesting and attracted my attention.
   (__ 1 __ 2 __ 3 __ 4 __ 5)

5. The technology platform (EthicsDebateAI) facilitated my participation effectively.
   (__ 1 __ 2 __ 3 __ 4 __ 5)

B. Learning & Skill Development:

1. This workshop increased my understanding of key ethical concepts in AI (e.g., bias, fairness, transparency, accountability). (__ 1 __ 2 __ 3 __ 4 __ 5)

2. The case studies helped me understand how ethical principles apply to real-world AI scenarios. (__ 1 __ 2 __ 3 __ 4 __ 5)

3. Participating in the debates improved my ability to analyze different perspectives on ethical issues. (__ 1 __ 2 __ 3 __ 4 __ 5)

4. This workshop improved my ability to articulate reasoned arguments about AI ethics. (__ 1 __ 2 __ 3 __ 4 __ 5)

5. The workshop activities encouraged me to think critically about the societal impact of AI. (__ 1 __ 2 __ 3 __ 4 __ 5)

C. Relevance & Career Preparedness:

1. The ethical dilemmas discussed felt relevant to my future career in computer science. (__ 1 __ 2 __ 3 __ 4 __ 5)

2. This workshop made me more aware of the ethical responsibilities of AI professionals. (__ 1 __ 2 __ 3 __ 4 __ 5)

3. I feel more prepared now to identify potential ethical issues in AI projects I might work on. (__ 1 __ 2 __ 3 __ 4 __ 5)

4. I feel more confident in my ability to discuss ethical concerns related to AI with future colleagues or managers. (__ 1 __ 2 __ 3 __ 4 __ 5)

5. This workshop provided valuable preparation for navigating ethical challenges in an AI-related career. (__ 1 __ 2 __ 3 __ 4 __ 5)

D. Overall Satisfaction:

1. Overall, I was satisfied with the "Career-Focused Workshop Series" on AI Ethics. (__ 1 __ 2 __ 3 __ 4 __ 5)

2. I would recommend this workshop (including the EthicsDebateAI platform) to other computer science students. (__ 1 __ 2 __ 3 __ 4 __ 5)

Part 2: Open-Ended Questions

Instructions: Please provide more detailed feedback on the spaces below.

1. What aspect(s) of the EthicsDebateAI platform (e.g., debate structure, case studies, resources) did you find most helpful for your learning about AI ethics? Why?

_____
_____

2. What aspect(s) of the platform or the workshop activities were the least helpful or could it be improved? Please provide specific suggestions.

_____
_____

3. Can you describe a specific skill or understanding related to AI ethics that you gained or significantly improved through this workshop?

_____
_____

4. Thinking about your future career, what specific ethical challenges related to AI do feel this workshop helped you prepare for? Are there any areas where you still feel unprepared?

_____

_____

5. Do you have any other comments, suggestions, or feedback about the workshop or the EthicsDebateAI platform?

_____

_____

Appendix D

Table D1
Descriptive Statistics for Individual Satisfaction Survey Items (N = 50)

Item Code	Item Description (Abbreviated)	Mean	Std. Deviation
A: Platform Usability & Engagement			
Sat_A1	Easy to navigate/use	4.18	0.69
Sat_A2	Instructions clear	4.20	0.61
Sat_A3	Debate format engaging	4.40	0.64
Sat_A4	Case studies interesting	4.48	0.58
Sat_A5	Platform facilitated participation	4.12	0.66
B: Learning & Skill Development			
Sat_B1	Increased understanding of concepts	4.54	0.50
Sat_B2	Understood application of principles	4.52	0.54
Sat_B3	Improved analysis of perspectives	4.32	0.71
Sat_B4	Improved articulation of arguments	4.18	0.66
Sat_B5	Encouraged critical thinking	4.56	0.54
C: Relevance & Career Preparedness			
Sat_C1	Dilemmas relevant to career	4.62	0.49
Sat_C2	More aware of AI professional responsibilities	4.72	0.45
Sat_C3	More prepared to identify ethical issues	4.18	0.72
Sat_C4	More confidently discussing ethical concerns	4.24	0.74
Sat_C5	Valuable prep for navigating challenges	4.46	0.58
D: Overall Satisfaction			
Sat_D1	Overall satisfied with workshop	4.52	0.54
Sat_D2	Would recommend workshop	4.44	0.68

Note. Scale: 1=Strongly Disagree to 5=Strongly Agree. SD = Standard Deviation.