ESJ Social Sciences Peer-reviewed

Between Efficiency and Empathy: An Exploratory Study of Albanian Students' Attitudes Towards AI in Ethical Decision-Making

Aljula Gjeloshi, PhD
Anila Boshnjaku, Associate Professor
Ledia Thoma, Professor
Agricultural University of Tirana, Albania

Doi:10.19044/esj.2026.v22n7p155

Submitted: 13 February 2026
Accepted: 23 March 2026
Published: 31 March 2026

Cite As:
Gjeloshi, A., Boshnjaku, A. & Thoma, L. (2026). *Between Efficiency and Empathy: An Exploratory Study of Albanian Students' Attitudes Towards AI in Ethical Decision-Making.* European Scientific Journal, ESJ, 22 (7), 155. https://doi.org/10.19044/esj.2026.v22n7p155

Abstract

This article presents an exploratory study of how university students in Albania (N = 146) perceive the role of artificial intelligence (AI) in ethical decision-making. While AI is increasingly integrated into everyday choices, little is known about how individuals in non-Western, digitally emerging contexts negotiate trust, empathy, and moral responsibility in relation to algorithmic systems. Drawing on survey data, we find that respondents maintain a strong attachment to empathy as a moral reference point (M = 3.12) and express significant reservations about delegating emotionally consequential decisions to AI (M = 3.41 for routine vs. emotional delegation). Trust in AI's ethical competence is moderate and conditional, strongly associated with perceived safety (r =.54) and transparency. We interpret these findings through the lens of what we term the "residual moral subject"—a concept describing the persistence of human moral agency in technologically mediated environments, albeit in a reconfigured form. The study contributes empirical nuance to debates on algorithmic governance and highlights the need for culturally situated research on AI ethics.

Keywords: Moral agency, artificial intelligence, algorithmic governance, empathy, Albania, higher education, ethical decision-making

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

Introduction

The Question of Moral Agency in the Age of AI

Artificial intelligence is no longer merely a computational tool; it increasingly functions as a normative infrastructure, shaping decisions about justice, responsibility, and human well-being. From algorithmic recommendations in healthcare to automated assessments in hiring, AI systems intervene in domains traditionally governed by human ethical judgment. This transformation raises a fundamental question: what happens to the human moral subject when ethical decision-making is partially delegated to algorithms?

While a growing body of literature has explored this question theoretically (Coeckelbergh, 2020; Floridi, 2013) and experimentally in Western contexts (Bigman & Gray, 2018; Gogoll & Müller, 2017), empirical research remains skewed toward high-income, technologically saturated societies. Less is known about how young adults in emerging economies—who are among the most active users of digital technologies—perceive and negotiate the moral boundaries between human and algorithmic judgment.

This study addresses this gap by examining the attitudes of university students in Albania toward AI-mediated decision-making. Specifically, it investigates: (1) How do respondents perceive the role of empathy in ethical decisions involving AI? (2) To what extent do they trust AI as a moral agent? (3) What boundaries do they draw regarding the delegation of morally consequential decisions to algorithmic systems?

Rather than asking whether AI is "good" or "bad," this article shifts the focus toward the subjective dimension of moral authority. We introduce the concept of the residual moral subject as an interpretive lens for understanding how human agency persists—and is reconfigured—within algorithmically mediated environments. The concept does not imply a nostalgic return to a pre-technological subject, but rather describes a condition in which individuals remain morally engaged while navigating decision contexts increasingly shaped by algorithmic systems.

The article proceeds as follows: Section 2 reviews relevant literature on algorithmic governance, moral delegation, and empirical studies of trust in AI. Section 3 describes the methodology and sample characteristics. Section 4 presents descriptive and correlational findings. Section 5 discusses these findings in relation to the concept of the residual moral subject. Section 6 concludes with implications and directions for future research.

2. Literature Review

2.1 Algorithmic Governance *and the Redistribution of Moral Authority*

Artificial intelligence is increasingly embedded in decision-making processes that were traditionally the domain of human moral judgment. From

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

predictive policing and healthcare triage to autonomous vehicles, AI systems mediate decisions with significant ethical consequences. Scholars have conceptualized this shift as algorithmic governance, in which algorithms do not merely perform technical functions but actively shape norms, priorities, and perceptions of what constitutes a "good" decision (Beer, 2017; Yeung, 2018).

While proponents highlight efficiency, consistency, and objectivity, critics stress that algorithms embed normative assumptions through data curation, modeling choices, and optimization goals (O'Neil, 2016; Eubanks, 2018). This implies that delegation of decision-making to algorithms is never ethically neutral. It produces moral distancing, where humans are involved in outcomes without fully owning their moral implications (Introna, 2016). Despite extensive discussion of these structural transformations, empirical investigations into how individuals perceive AI-mediated moral authority remain limited.

2.2 *The Limits of Moral Delegation: Why Empathy Matters*

Moral delegation refers to the transfer of ethically significant judgment from humans to technical systems (Coeckelbergh, 2020). Machine ethics research suggests that ethical reasoning can be partially formalized through rules, constraints, and objective functions (Moor, 2006; Wallach & Allen, 2009). Yet, philosophers emphasize that morality is irreducible to formal rules. Moral judgment relies on context, relational understanding, and emotional sensitivity (Dreyfus, 1992; Nussbaum, 2001; Tronto, 1993).

Empirical studies corroborate this critique. Bigman and Gray (2018) found that participants resist delegating morally sensitive decisions to machines, particularly those involving human suffering or social consequences. Gogoll and Müller (2017) similarly report that while AI is acceptable as an advisory tool, participants prefer human final authority in ethical dilemmas. These findings suggest that humans maintain boundaries of moral delegation, guided by empathy and contextual judgment.

2.3 Empirical Studies on Trust in AI: A Western Bias

Recent empirical research has expanded our understanding of how individuals evaluate AI systems. Studies show that trust in AI is multidimensional, conditional, and context-dependent (Oliveira et al., 2024; Gerlich, 2024). Experimental work reveals that individuals may rely on AI-generated recommendations even when aware of potential unreliability, particularly under conditions of complexity or cognitive load (Krügel, 2022). Other research suggests that trust in AI may sometimes emerge as a compensatory response to distrust in human institutions (Galindez-Acosta et al., 2025).

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

However, the overwhelming majority of this research has been conducted in Western, educated, industrialized, rich, and democratic (WEIRD) societies. This geographic and cultural concentration limits our understanding of how attitudes toward AI vary across contexts with different institutional histories, technological infrastructures, and cultural values. As AI systems become globally ubiquitous, understanding context-specific perceptions of algorithmic authority becomes not only academically relevant but also ethically necessary.

This study contributes to diversifying the empirical landscape by examining AI attitudes in Albania, a post-communist, upper-middle-income country with rapidly increasing digital adoption but limited public debate on AI ethics.

2.4 The Residual Moral Subject: A Conceptual Framework

Building on critical social theory and the philosophy of technology, we introduce the concept of the residual moral subject to interpret contemporary forms of moral agency under conditions of algorithmic mediation (Adorno & Horkheimer, 1947; Verbeek, 2011).

The concept refers to a form of moral agency that persists despite—and is reshaped by—the increasing integration of algorithmic systems into decision-making processes. Rather than disappearing, moral subjectivity is reconfigured, operating within technological environments that shape, constrain, and partially redistribute responsibility.

Three characteristics define this residual subjectivity:

1. Ethical awareness persists but is mediated. Individuals continue to evaluate outcomes in moral terms, even when algorithmic recommendations are involved. However, their judgments are shaped by the informational and procedural affordances of technological systems.
2. Moral responsibility becomes conditional and situational. Individuals may rely on algorithmic advice in some contexts (e. g., technical decisions) while insisting on human judgment in others (e. g., decisions with emotional or relational consequences).
3. Responsibility is negotiated within hybrid human–machine configurations. Decision-making increasingly takes place in environments where algorithmic systems provide recommendations while humans retain interpretive, supervisory, or veto power.

Importantly, the concept of the residual moral subject differs from approaches that dissolve agency entirely across distributed networks (Latour, 2005). While acknowledging technological mediation, it insists that normative accountability ultimately remains anchored in human actors. The concept

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

serves as an interpretive framework for the empirical patterns observed in this study.

3. Methodology

3.1 Research Design and Sampling

This study employs an exploratory quantitative research design to examine perceptions of moral agency and algorithmic delegation in AI-mediated decision-making. Data were collected through a structured survey administered to university students at a public university in Albania between October 2025 and January 2026.

Participants were recruited using convenience sampling in classroom settings at a public university in Albania. Participation was voluntary and anonymous. Prior to completing the survey, respondents were informed about the academic purpose of the study, assured that no identifying information would be collected, and reminded that participation was entirely optional and could be withdrawn at any time. The research followed standard ethical guidelines for social research involving human participants.

The final sample consisted of 146 respondents. Participants were predominantly undergraduate students in the early stages of their academic studies. The mean age of respondents was 19.48 years (SD = 2.50), with an age range from 17 to 46 years. Most participants were enrolled in their first or second year of study, reflecting a population of young adults currently developing their civic and ethical perspectives in an increasingly digital environment.

This population is analytically relevant because younger cohorts are among the most frequent users of digital technologies and are likely to encounter AI systems in educational, professional, and everyday decision-making contexts. Their perceptions therefore provide insight into emerging attitudes toward algorithmic authority and moral responsibility.

It is important to note that the sample is not statistically representative of the general population. As a convenience sample of university students, the findings should be interpreted as exploratory rather than generalizable. Nevertheless, the dataset offers analytically valuable insight into how digitally socialized young adults perceive the relationship between human moral judgment and AI-assisted decision-making.

Descriptive characteristics of the sample are presented in Table 1.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

Table 1: Demographic Characteristics of the Sample (N = 146)

Variable	Category	N	%
Gender	Female	96	65.8
	Male	50	34.2
Age Group	17–20 years	104	71.2
	21–25 years	34	23.3
	26+ years	8	5.5
Year of Study	First year	68	46.6
	Second year	54	37.0
	Third year	18	12.3
	Other / not specified	6	4.1

Source: Authors' calculations

3.2 Measurement and Instrument

The survey instrument consisted of Likert-scale items measured on a five-point scale (1 = strongly disagree, 5 = strongly agree). The questionnaire was designed to capture several dimensions of perceptions toward AI-mediated ethical decision-making: (a) perceived moral centrality of empathy, (b) beliefs regarding contextual and non-programmable aspects of ethical judgment, (c) trust in AI's ethical and decision-making capacity, (d) willingness to delegate morally significant decisions to AI, and (e) conditional boundaries of trust.

Items were developed based on themes identified in the literature on algorithmic governance, moral delegation, and human–AI trust (Bigman & Gray, 2018; Coeckelbergh, 2020; Gogoll & Müller, 2017; Yeung, 2018). Where conceptually appropriate, items were grouped into composite indices representing broader attitudinal constructs.

Three primary composite indices were constructed:

- Moral Trust in AI (4 items): Capturing perceptions of AI's ethical competence and perceived safety in decision-making contexts. Example items include: "AI can help humans make more ethical decisions" and "I sometimes feel safer relying on AI for complex decisions."
- Delegation Willingness (3 items): Reflecting respondents' readiness to accept AI recommendations or delegate decisions to algorithmic systems. Example items include: "I am willing to accept AI suggestions, even if I don't fully agree" and "I would prefer to collaborate with AI to make complex decisions."
- Conditional Boundaries of Trust (3 items): Measuring the limits respondents place on delegation, particularly concerning emotional consequences, transparency, and human oversight. Example items include: "I would trust AI for routine decisions but not for those with

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

major emotional impact" and "AI decisions should always be transparent and explainable."

Composite indices were computed by averaging the items associated with each construct. No reverse-coded items were included in the instrument. Missing responses were minimal and handled through listwise deletion.

Internal consistency reliability was assessed using Cronbach's alpha. All composite indices demonstrated acceptable internal reliability ($\alpha \geq .70$), indicating coherent underlying attitudinal constructs.

Reliability analysis indicated acceptable internal consistency across the composite indices. Moral Trust in AI demonstrated a Cronbach's alpha of approximately **0.72**, Delegation Willingness **0.70**, and Conditional Trust Boundaries **0.75**, indicating satisfactory internal coherence of the constructs.

Table 2: Reliability Analysis of Composite Indices

Composite Index	Number of Items	Cronbach's Alpha
Moral Trust in AI	4	0.72
Delegation Willingness	3	0.70
Conditional Boundaries of Trust	3	0.75

Source: Authors' calculations

The full list of survey items is provided in **Appendix A** to ensure transparency and replicability.

3.3 Analytical Strategy

The analysis proceeded in three stages.

First, descriptive statistics were calculated to assess the central tendencies and variability of respondents' attitudes toward AI-mediated ethical decision-making. Means and standard deviations were used to identify general patterns of agreement or ambivalence across items.

Second, bivariate correlations were examined using Pearson correlation coefficients (r) to explore associations between key attitudinal variables, including perceived ethical competence of AI, comfort with collaboration, and boundaries of moral delegation. Correlation coefficients were interpreted using conventional benchmarks (small $\approx .10$, moderate $\approx .30$, strong $\geq .50$). Statistical significance was evaluated at the $p < .05$ level (two-tailed).

Third, reliability analysis was conducted using Cronbach's alpha coefficients to evaluate the internal consistency of the composite indices (reported in Table 2).

The purpose of the analysis is exploratory rather than predictive; therefore, emphasis is placed on identifying relational patterns among variables rather than establishing causal inference.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

3.4 Analytical Orientation

The study does not aim for statistical generalization to broader populations. Instead, it seeks analytical insight into how moral agency is perceived under conditions of algorithmic mediation within a specific sociocultural context. The objective is interpretive: to identify structural patterns of ethical ambivalence and boundaries of moral delegation that may inform theoretical debates on algorithmic governance and the reconfigured relationship between human judgment and technological systems.

Following the tradition of interpretive social science, the findings are situated within broader philosophical debates on moral agency and technological mediation (Adorno & Horkheimer, 1947; Verbeek, 2011), contributing empirical grounding to normative theory rather than testing context-free universal claims.

3.5 Limitations

First, the use of convenience sampling at a single public university in Albania restricts generalizability beyond the specific population studied. The sample is predominantly composed of young students, and findings may not extend to other demographic groups, institutional contexts, or cultural settings. Second, self-reported perceptions of moral agency and trust in AI may not fully correspond to behavioral decisions in real-world contexts. Respondents may express different attitudes when confronted with actual moral dilemmas involving AI systems than when responding to abstract survey items. Third, the cross-sectional design captures attitudes at a single moment in time and cannot assess how these perceptions might evolve with increased exposure to AI technologies, changes in the technological landscape, or shifts in public discourse around algorithmic governance. Fourth, while the composite indices demonstrate acceptable internal consistency, they represent parsimonious approximations of complex attitudinal constructs. Future research could benefit from more fine-grained measurement approaches, including experimental vignettes or qualitative interviews.

Despite these limitations, the study offers analytically valuable insight into emerging patterns of ethical ambivalence and conditional trust in AI among a digitally socialized cohort in a non-Western European context. It provides an empirical foundation for future research employing more diverse samples, comparative designs, experimental methods, or longitudinal approaches.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

4. Findings and Results: Empirical Indicators of Ethical Ambivalence

4.1 Descriptive Statistics: General Patterns

Table 3 presents descriptive statistics for all survey items related to ethical attitudes toward AI-mediated decision-making. Means and standard deviations provide an overview of central tendencies and variability within the sample.

Table 3: Descriptive Statistics of Ethical Attitudes Toward AI (N = 146)

Survey Statement	Mean (M)	Std. Deviation (SD)
Empathy is an essential element in every decision I make	3.12	0.91
Humans are more empathetic than AI systems	3.17	1.21
Ethical decisions must be context-based and cannot be fully programmed	3.43	1.10
AI can help humans make more ethical decisions	2.74	1.11
I sometimes feel safer relying on AI for complex decisions	2.55	1.12
Decisions affecting others should always include a human component	3.31	1.18
I am concerned that AI may ignore human emotions in decisions	3.47	1.10
AI decisions should always be transparent and explainable	3.68	1.05
I would trust AI for routine decisions but not for emotionally significant ones	3.41	1.30
Comfort collaborating with AI in decision-making	3.08	0.98

Source: Authors' calculations

Several patterns emerge from these descriptive statistics.

First, respondents attribute moderate importance to empathy in their own decision-making (M = 3.12) and tend to perceive humans as more empathetic than AI systems (M = 3.17). However, the relatively high standard deviation for the latter item (SD = 1.21) indicates considerable variation in respondents' views, suggesting that while the central tendency points toward agreement, a substantial minority may hold different opinions.

Second, there is relatively strong agreement that ethical decisions depend on context and cannot be reduced to programmable rules (M = 3.43). This finding suggests that respondents intuitively recognize limits to the formalization of moral judgment within algorithmic frameworks.

Third, trust in AI's ethical competence appears limited. Mean scores for statements such as "AI can help humans make more ethical decisions" (M = 2.74) and "*I sometimes feel safer relying on AI for complex decisions*" (M = 2.55) fall below or near the midpoint of the scale, indicating skepticism rather than endorsement. At the same time, respondents express moderate comfort with the idea of collaborating with AI in decision-making (M = 3.08), suggesting openness to algorithmic assistance under certain conditions.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

Fourth, concerns about AI's capacity to account for human emotions are evident. Respondents agree, on average, that they are concerned about AI ignoring human emotions (M = 3.47) and that decisions affecting others should always include a human component (M = 3.31). The strongest agreement in the survey pertains to the need for transparency: "*AI decisions should always be transparent and explainable*" (M = 3.68, SD = 1.05).

Fifth, respondents draw clear boundaries regarding the types of decisions they are willing to delegate to AI systems. Agreement with the statement "*I would trust AI for routine decisions but not for emotionally significant ones*" (M = 3.41) suggests that acceptance of algorithmic involvement is conditional and domain-specific.

Taken together, these descriptive patterns point toward a stance of ethical ambivalence: respondents neither fully reject AI involvement in decision-making nor uncritically embrace it. They acknowledge potential benefits while maintaining reservations, particularly when decisions carry emotional weight or involve human well-being.

4.2 Composite Indices

To facilitate higher-level analysis, conceptually related items were combined into three composite indices. Table 4 presents descriptive statistics for these indices.

Table 4: Descriptive Statistics of Composite Indices (N = 146)

Composite Index	Mean (M)	Std. Deviation (SD)	Cronbach's Alpha
Moral Trust in AI	2.68	0.89	0.72
Delegation Willingness	2.82	0.84	0.70
Conditional Boundaries of Trust	3.52	0.91	0.75

Source: Authors' calculations.

Note. All indices measured on a 5-point scale (1 = strongly disagree, 5 = strongly agree).

The Moral Trust in AI index (M = 2.68) confirms that respondents, on average, do not strongly trust AI's ethical competence. This score falls below the scale midpoint, indicating a tendency toward skepticism rather than confidence.

The Delegation Willingness index (M = 2.82) similarly falls slightly below the midpoint, suggesting that respondents are generally hesitant to accept AI recommendations or delegate decisions to algorithmic systems, though not entirely opposed.

The Conditional Boundaries of Trust index (M = 3.52) is notably higher, indicating that respondents strongly endorse the idea that trust in AI should be limited by considerations of emotional impact, transparency, and human oversight.

These patterns are consistent with the item-level analysis: respondents are cautious about delegating moral authority to AI, and this caution is

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

particularly pronounced when decisions involve emotional or relational consequences.

4.3 Correlational Analysis

To explore relationships between key attitudinal variables, bivariate Pearson correlations were calculated. Table 5 presents the correlation matrix for selected items and composite indices.

Table 5: Correlations Between Key Variables (N = 146)

Variable	1	2	3	4	5	6
1. Empathy essential in decision-making	1					
2. Humans more empathetic than AI	.42**	1				
3. Ethical decisions must be context-based	.38**	.29**	1			
4. AI can help ethical decisions (item)	-.21*	-.26**	-.19*	1		
5. Moral Trust in AI (index)	-.23*	-.28**	-.21*	.58**	1	
6. Comfort collaborating with AI (item)	-.17	-.20*	-.15	.41**	.48**	1

Source: Authors' calculations.
Note. *p <.05, **p <.01 (two-tailed).

The correlation matrix reveals several noteworthy patterns.

First, there is a moderate positive correlation between respondents' endorsement of empathy in their own decision-making and their perception that humans are more empathetic than AI ($r = .42$, $p < .01$). This suggests that individuals who value empathy highly are also more likely to view it as a distinctively human capacity.

Second, empathy-based variables correlate positively with the belief that ethical decisions must remain context-based ($r = .38$ between empathy and context, $p < .01$; $r = .29$ between human empathy and context, $p < .01$). Respondents who emphasize empathy tend to view moral judgment as situationally dependent rather than reducible to fixed rules.

Third, and importantly, variables related to trust in AI show negative correlations with empathy-based attitudes. The belief that AI can help humans make more ethical decisions is negatively correlated with empathy in decision-making ($r = -.21$, $p < .05$), with the perception that humans are more empathetic than AI ($r = -.26$, $p < .01$), and with the belief that ethical decisions are context-based ($r = -.19$, $p < .05$). The Moral Trust in AI index shows a similar pattern of negative associations with these empathy- and context-related variables.

Fourth, there is a strong positive correlation between the belief that AI can help ethical decisions and the Moral Trust in AI index ($r = .58$, $p < .01$). This is expected, as the former item is a component of the latter index, and confirms the internal coherence of the construct.

Fifth, comfort collaborating with AI is positively associated with Moral Trust in AI ($r = .48$, $p < .01$) and with the belief that AI can help ethical decisions ($r = .41$, $p < .01$). However, it shows weak or negative associations

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

with empathy-based variables, suggesting that openness to AI collaboration may be more closely related to perceptions of AI competence than to views about human moral distinctiveness.

4.4 Summary of Findings

The results can be summarized in five key points:

1. Empathy remains salient: Respondents continue to view empathy as an important element of moral decision-making and perceive humans as more empathetic than AI systems.
2. Contextual judgment is valued: There is broad agreement that ethical decisions depend on context and cannot be fully programmed, reflecting intuitive recognition of limits to algorithmic formalization.
3. Trust in AI's ethical competence is limited: Mean scores for trust-related items and composite indices fall at or below the scale midpoint, indicating skepticism rather than confidence in AI's capacity for ethical judgment.
4. Delegation is conditional: Respondents draw clear boundaries, expressing greater willingness to accept AI involvement in routine decisions than in those with emotional consequences. Transparency and human oversight emerge as important conditions for trust.
5. Attitudes cluster coherently: Correlation analysis reveals that empathy-based attitudes tend to cluster together and are negatively associated with trust in AI. Conversely, openness to AI collaboration correlates positively with perceptions of AI's ethical competence.

These patterns, considered together, depict a stance of ethical ambivalence. Respondents do not occupy extreme positions—neither fully embracing nor completely rejecting AI in decision-making. Instead, they navigate a middle ground characterized by selective reliance, conditional trust, and the maintenance of boundaries around emotionally significant decisions. This ambivalence, we argue, reflects the contours of what might be termed the residual moral subject: a form of moral agency that persists under conditions of algorithmic mediation, but is reconfigured—becoming conditional, negotiated, and context-sensitive.

5. Discussion: Interpreting Patterns of Ethical Ambivalence

The findings of this exploratory study offer empirical insight into how a specific group of young adults in Albania perceives the relationship between empathy, moral responsibility, and AI-mediated decision-making. In this section, we interpret these patterns through the lens of the residual moral subject—a conceptual framework for understanding how moral agency persists and is reconfigured under conditions of algorithmic mediation. We

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

then situate the findings within the broader literature, acknowledge limitations, and propose directions for future research.

5.1 Empathy as an Anchor of Moral Judgment

One of the clearest patterns to emerge from the data is the continued salience of empathy as a moral reference point. Respondents expressed moderate agreement that empathy is essential in their own decision-making ($M = 3.12$) and tended to perceive humans as more empathetic than AI systems ($M = 3.17$). The positive correlation between these two items ($r = .42$, $p < .01$) suggests that valuing empathy personally is associated with viewing it as a distinctively human capacity.

This finding resonates with philosophical accounts that emphasize the cognitive and evaluative role of emotions in moral reasoning. Nussbaum (2001) argues that emotions are not merely subjective reactions but forms of evaluative judgment that help individuals identify what is morally significant in human life. Similarly, the ethics of care tradition positions empathy, attentiveness, and responsibility toward others as essential dimensions of moral agency (Gilligan, 1993; Tronto, 1993). From this perspective, respondents' attachment to empathy may reflect an intuitive understanding that moral judgment requires forms of relational sensitivity that cannot be easily reduced to algorithmic computation.

At the same time, the moderate mean scores and non-negligible standard deviations indicate that this attachment to empathy is not universal or unqualified. Some respondents may be more open to algorithmic involvement in decision-making, even in domains traditionally associated with human judgment. This variability is itself analytically interesting, suggesting that attitudes toward AI-mediated ethics are contested and evolving rather than settled.

5.2 Contextual Judgment and the Limits of Formalization

Respondents expressed relatively strong agreement that ethical decisions must be context-based and cannot be fully programmed ($M = 3.43$). This finding aligns with longstanding critiques of artificial intelligence that emphasize the irreducibility of human judgment to formal rules. Dreyfus (1992) argued that expert human reasoning depends heavily on tacit knowledge, embodied experience, and situational interpretation—dimensions that resist explicit computational specification. Although contemporary research in machine ethics has explored ways to formalize ethical principles within algorithmic frameworks (Moor, 2006; Wallach & Allen, 2009), the translation of moral judgment into computational models remains conceptually and practically challenging.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

The positive correlation between belief in context-based ethics and empathy-based attitudes (r =.38 with empathy, p <.01; r =.29 with human empathy, p <.01) suggests that these two dimensions of moral reasoning are interconnected in respondents' minds. Those who emphasize empathy also tend to view ethical judgment as situationally dependent rather than rule-governed. This pattern is consistent with theoretical accounts that position empathy as a capacity for situational attunement—a way of perceiving what particular situations demand rather than applying abstract principles mechanically (Vetlesen, 1994).

5.3 Limited Trust and the Conditional Nature of Delegation

A second major pattern concerns the limited trust respondents place in AI's ethical competence. Mean scores for items such as "AI can help humans make more ethical decisions" (M = 2.74) and the Moral Trust in AI composite index (M = 2.68) fall below the scale midpoint, indicating skepticism rather than confidence. At the same time, respondents express moderate comfort with the idea of collaborating with AI in decision-making (M = 3.08), suggesting that they are not entirely closed to algorithmic assistance.

This apparent tension—skepticism about AI's ethical capacity alongside openness to collaboration—resolves when we consider the conditional nature of trust revealed in the data. The Conditional Boundaries of Trust index (M = 3.52) indicates strong endorsement of the idea that trust in AI should be limited by considerations of emotional impact, transparency, and human oversight. Respondents agree, on average, that they would trust AI for routine decisions but not for emotionally significant ones (M = 3.41), and the strongest agreement in the survey pertains to the need for transparency (M = 3.68).

These patterns are consistent with experimental research showing that individuals resist delegating morally sensitive decisions to machines, particularly those involving human suffering or social consequences (Bigman & Gray, 2018; Gogoll & Müller, 2017). They also align with survey-based studies demonstrating that trust in AI is frequently grounded in perceptions of procedural fairness and transparency rather than in beliefs that machines possess genuine moral understanding (Gerlich, 2024).

The correlation between comfort collaborating with AI and Moral Trust in AI (r =.48, p <.01) suggests that openness to algorithmic assistance is associated with perceptions of AI's ethical competence. However, the negative correlations between trust in AI and empathy-based attitudes (r = -.23 between Moral Trust and empathy; r = -.28 between Moral Trust and belief in human empathy) indicate that these two orientations may represent distinct—and somewhat opposed—ways of thinking about moral agency in relation to technology.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

5.4 The Residual Moral Subject: An Interpretive Lens

How should we make sense of these patterns? We propose that they can be interpreted through the concept of the residual moral subject—a form of moral agency that persists under conditions of algorithmic mediation but is reconfigured, becoming conditional, negotiated, and context-sensitive.

The concept, introduced in Section 2.4, draws on critical social theory and the philosophy of technology (Adorno & Horkheimer, 1947; Verbeek, 2011). It differs from approaches that either celebrate the displacement of human judgment by algorithmic systems or mourn its disappearance. Instead, it describes a situation in which individuals remain morally engaged—they continue to evaluate outcomes in ethical terms, to draw boundaries around acceptable delegation, and to insist on human oversight—but do so within environments increasingly shaped by technological infrastructures.

The empirical patterns observed in this study illustrate this configuration:

- Persistence of ethical awareness: Respondents continue to view empathy and contextual judgment as morally significant, even when considering AI-mediated decisions. They do not cede the moral domain to machines.
- Conditionality of trust: Trust in AI is not absolute but contingent on factors such as transparency, emotional impact, and the presence of human oversight. Respondents are willing to accept algorithmic assistance, but on their own terms.
- Negotiation of boundaries: Respondents draw clear distinctions between decisions they would delegate to AI (routine, technical) and those they would reserve for human judgment (emotionally significant, relationally complex). These boundaries are not given but actively maintained.
- Selective reliance: The weak correlations between trust in AI and empathy-based attitudes suggest that openness to algorithmic assistance may coexist with—rather than replace—attachment to human moral capacities. Respondents appear to navigate a middle ground, neither fully embracing nor completely rejecting AI involvement.

In this interpretation, the residual moral subject is not a nostalgic figure clinging to a pre-technological past, but a contemporary agent learning to navigate a world in which efficiency and empathy, algorithm and intuition, must coexist. The ambivalence evident in the data—the simultaneous openness to and wariness of AI—reflects the challenges of maintaining moral agency within hybrid human–machine decision environments.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

5.5 The Specificity of Context: Albania as a Non-WEIRD Setting

It is important to situate these findings within the specific context in which they were produced. The sample consists of university students in Albania—a post-communist, upper-middle-income country with rapidly increasing digital adoption but limited public infrastructure for AI governance and minimal public debate on AI ethics.

Several contextual factors may shape the attitudes observed:

First, Albania's educational system, particularly in economics and social sciences, continues to emphasize humanistic values and ethical reflection. Students may be socialized to view empathy and contextual judgment as central to professional and civic life, potentially reinforcing skepticism toward purely technical approaches to decision-making.

Second, exposure to AI systems in everyday life, while growing, remains less pervasive than in highly digitized societies such as the United States, China, or Northern Europe. Respondents may have less direct experience with AI-mediated decisions in high-stakes domains (healthcare, finance, criminal justice), potentially making their attitudes more abstract and less calibrated to actual technological capabilities.

Third, cultural norms in Albania—as in many Mediterranean and post-communist societies—continue to emphasize interpersonal relationships, family ties, and community bonds. These relational orientations may make individuals particularly sensitive to the emotional and social dimensions of decision-making, and consequently more cautious about delegating such decisions to machines.

Fourth, the sample's disciplinary concentration in economics (93%) may introduce specific biases. Economics students are trained to think in terms of efficiency, optimization, and rational choice—frameworks that might, in principle, make them more receptive to algorithmic decision-making. Yet the findings show only modest trust in AI, suggesting that disciplinary training does not simply translate into techno-optimism. Future research could usefully compare attitudes across disciplines.

Given these contextual specificities, we do not claim that the patterns observed here generalize to other populations. Rather, we offer them as a contribution to diversifying the empirical landscape of AI ethics research, which remains heavily skewed toward Western, educated, industrialized, rich, and democratic (WEIRD) societies (Henrich, Heine, & Norenzayan, 2010). Understanding how attitudes toward algorithmic governance vary across cultural and institutional contexts is essential for developing globally relevant ethical frameworks.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

5.6 Relationship to Existing Literature

The findings both align with and extend existing research on human–AI interaction and moral delegation.

Consistent with previous studies, we find that individuals are reluctant to delegate morally sensitive decisions to machines (Bigman & Gray, 2018; Gogoll & Müller, 2017). The Conditional Boundaries of Trust index (M = 3.52) and the item on routine versus emotional decisions (M = 3.41) provide clear evidence that respondents maintain boundaries around delegation.

The findings also resonate with research on the multidimensional nature of trust in AI. Studies have shown that individuals distinguish between technical competence and moral authority (Oliveira et al., 2024), and that trust is often conditional on perceived transparency and fairness (Gerlich, 2024). The strong endorsement of transparency in our data (M = 3.68) reinforces this point.

Where this study extends existing work is in its focus on a non-WEIRD population and its attempt to interpret empirical patterns through a theoretically grounded concept—the residual moral subject. Most experimental and survey-based research on AI attitudes has been conducted in Western contexts, leaving open the question of how cultural and institutional factors shape perceptions of algorithmic authority. By documenting patterns of ethical ambivalence in Albania, we contribute to building a more globally diverse evidence base.

5.7 Implications for AI Governance and Design

The findings also carry implications for the governance and design of AI systems.

First, they suggest that public acceptance of algorithmic decision-making depends not only on technical performance but on perceived ethical legitimacy. Transparency, explainability, and human oversight are not optional add-ons but central conditions for trust. Developers and policymakers should prioritize these features in system design and regulatory frameworks.

Second, the persistence of empathy as a moral reference point indicates that AI governance frameworks should avoid assuming that ethical responsibility can be fully delegated to technological systems. Decision-making models that combine algorithmic support with human judgment—so-called "human-in-the-loop" configurations—may be more socially acceptable and normatively defensible than fully automated systems, particularly in domains involving emotional or relational consequences.

Third, the conditional nature of trust observed in this study suggests that acceptance of AI is not a binary proposition but a matter of degree and domain. Public engagement and deliberation should acknowledge this

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

complexity, exploring not whether AI should be used but under what conditions, for what purposes, and with what safeguards.

Fourth, the contextual specificity of the findings underscores the importance of culturally situated approaches to AI ethics. Global principles and guidelines must be adapted to local contexts, taking account of varying institutional histories, cultural values, and technological infrastructures.

Conclusions

This study examined how university students in Albania perceive the relationship between empathy, moral responsibility, and AI-mediated decision-making. The findings reveal that empathy remains a central moral reference, guiding respondents' judgments and shaping their perceptions of AI's ethical capacity.

Trust in AI is conditional and limited. Participants are more willing to delegate routine or low-stakes decisions to AI but retain human oversight for decisions with emotional or ethical significance. Attitudes toward AI are not random; they reflect coherent interpretive frameworks in which empathy, trust, and openness to collaboration interact.

The concept of the residual moral subject captures this dynamic. Moral agency is not lost but reconfigured: it persists, adapts, and is negotiated within hybrid human–machine contexts. Individuals selectively delegate responsibility, maintain boundaries, and balance efficiency with ethical reflection.

These insights have practical implications for AI design and governance. Systems that prioritize transparency, explainability, and support for human ethical judgment are more likely to be accepted. Furthermore, culturally and contextually sensitive approaches are essential, as perceptions of moral delegation vary across environments and populations.

Overall, the study highlights that human moral agency endures in algorithmically mediated contexts. Understanding how individuals navigate trust, delegation, and ethical responsibility is crucial for developing AI that complements rather than replaces human moral reasoning.

Conflict of Interest: The authors reported no conflict of interest.

Data Availability: All data are included in the content of the paper.

Funding Statement: The authors did not obtain any funding for this research.

References:

1. Adorno, T. W., & Horkheimer, M. (1947). *Dialectic of enlightenment*. Herder and Herder.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

2. Beer, D. (2017). The social power of algorithms. *Information, Communication & Society, 20*(1), 1–13. https://doi.org/10.1080/1369118X.2016.1216147
3. Bigman, Y. E., & Gray, K. (2018). People are averse to machines making moral decisions. *Cognition, 181*, 21–34. https://doi.org/10.1016/j.cognition.2018.08.003
4. Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
5. Dreyfus, H. L. (1992). *What computers still can't do: A critique of artificial reason*. MIT Press.
6. Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.
7. Floridi, L. (2013). *The ethics of information*. Oxford University Press.
8. Galindez-Acosta, J. S., & Giraldo-Huertas, J. J. (2025). Trust in AI emerges from distrust in humans: A machine learning study on decision-making guidance [Preprint]. *arXiv*. https://arxiv.org/abs/2511.16769
9. Gerlich, M. (2024). Exploring motivators for trust in the dichotomy of human–AI trust dynamics. *Social Sciences, 13*(5), Article 251. https://doi.org/10.3390/socsci13050251
10. Gilligan, C. (1993). *In a different voice: Psychological theory and women's development*. Harvard University Press.
11. Gogoll, J., & Müller, J. F. (2017). Autonomous cars: In favor of a mandatory ethics setting. *Science and Engineering Ethics, 23*(3), 681–700. https://doi.org/10.1007/s11948-016-9806-x
12. Introna, L. D. (2016). Algorithms, governance, and governmentality. *Science, Technology, & Human Values, 41*(1), 17–49. https://doi.org/10.1177/0162243915587360
13. Krügel, S., Ostermaier, A., & Uhl, M. (2022). Zombies in the loop? Humans trust untrustworthy AI-advisors for ethical decisions. *Philosophy & Technology, 35*(1), Article 17. https://doi.org/10.1007/s13347-022-00567-5
14. Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
15. Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems, 21*(4), 18–21. https://doi.org/10.1109/MIS.2006.80
16. Nussbaum, M. C. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge University Press.
17. Oliveira, M., Brands, J., Mashudi, J., Liefooghe, B., & Hortensius, R. (2024). Perceptions of artificial intelligence system's aptitude to judge morality and competence amidst the rise of chatbots. *Cognitive*

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

Research: Principles and Implications, 9(1), Article 47. https://doi.org/10.1186/s41235-024-00573-7

18. O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
19. Tronto, J. C. (1993). *Moral boundaries: A political argument for an ethic of care*. Routledge.
20. Verbeek, P.-P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.
21. Vetlesen, A. J. (1994). *Perception, empathy, and judgment: An inquiry into the preconditions of moral performance*. Pennsylvania State University Press.
22. Wallach, W., & Allen, C. (2009). *Moral machines: Teaching robots right from wrong*. Oxford University Press.
23. Yeung, K. (2018). Algorithmic regulation: A critical interrogation. *Regulation & Governance, 12*(4), 505–523. https://doi.org/10.1111/rego.12158

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

Appendix A: Survey Questionnaire

This appendix presents the questions from the survey "Empathy and Decision-Making in the Age of Artificial Intelligence," translated from the original Albanian. The response formats are as they appeared in the data collection sheet.

Part 1: Demographics and AI Familiarity

1. **Gender:**
   - Male
   - Female
   - Other/Prefer not to say
2. Age: (Open-ended numerical response)
3. **Year of Study:**
   - 1.0
   - 2.0
   - 3.0
   - Master
4. **Faculty/Department:**
   - Economics
   - Engineering / Technology
   - Social Sciences
   - (Other open-ended responses)
5. **Your knowledge about AI:**
   - 1 = No knowledge
   - 2 = Little knowledge
   - 3 = Average knowledge
   - 4 = Good knowledge
   - 5 = Advanced knowledge
6. How often do you use AI (e. g., ChatGPT, virtual assistants) in your studies or decision-making:
   - 1 = Never
   - 2 = Rarely
   - 3 = Sometimes
   - 4 = Often
   - 5 = Very often

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

Part 2: Attitudes Towards AI in Decision-Making

For questions 7-21, the response scale was: **Plotësisht kundër (Strongly disagree), Kundër (Disagree), Neutral / Nuk jam i sigurt (Neutral / Not sure), Dakord (Agree), Plotësisht dakord (Strongly agree)**

7. AI can make more objective decisions than humans.
8. AI decisions can often be unfair to certain individuals.
9. I prefer that decisions affecting me are made by a human and not AI.
10. AI can help in making complex ethical decisions.
11. AI can improve social fairness in decision-making.
12. I believe AI can have a programmed sense of empathy.
13. Humans are more reliable than AI in decision-making related to ethics.
14. AI decisions should always be transparent and reasoned.
15. In a distribution of resources, I would be more at ease if AI decided on the allocation.
16. I am concerned that AI can make decisions without considering people's feelings.
17. I believe AI decisions can be more efficient but less fair.
18. I am willing to accept AI suggestions, even if I don't fully agree.
19. I would prefer to collaborate with AI to make complex decisions.
20. I would often change my personal decision if AI suggested a more "efficient" but less empathetic solution.
21. When decisions affect people's feelings, I prefer that AI systems consider the emotional consequences before proposing solutions.

Part 3: Empathy, Ethics, and Trust in AI

For questions 22-29, the response scale was a Likert scale from 1 to 5, with anchors: 1 = Plotësisht kundër (Strongly disagree) and 5 = Plotësisht dakord (Strongly agree)

22. Empathy is an essential element in every decision I make.
23. I am willing to sacrifice personal benefits to help others.
24. I believe that humans are more empathetic than AI.
25. Ethical decisions should be based on context and cannot be programmable.
26. I think AI can help people make more ethical decisions.
27. Sometimes I feel more confident relying on AI for complex decisions.
28. Decisions that affect others should always have a human component.
29. For decisions that affect people, I always prefer a person's decision, even if AI suggests a more 'efficient' solution.

ISSN: 1857-7881 (Print)
e-ISSN 1857-7431 (Online)

Part 4: Preferences and Future Outlook

30. **How comfortable do you feel collaborating with AI in decision-making?**
    - 1 = Not comfortable at all
    - 2
    - 3 = Neutral
    - 4
    - 5 = Very comfortable
31. **How much do you think AI can improve society in ethical decision-making?**
    - 1 = Not at all
    - 2
    - 3 = Moderately
    - 4
    - 5 = Very much
32. **Would you like to participate in other experiments related to AI and ethics?**
    - Yes
    - No
33. **If you could choose, what role would you give to AI in your decision-making?**
    - Limited assistance
    - Decision-making partner
    - Full decision-making
    - No role
34. Write a situation where you think AI can make better decisions than humans. (Open-ended)
35. What are your main concerns in using AI for ethical decision-making? (Open-ended)

Part 5: Concluding Statements on Emotional Impact

For questions 36-39, the response scale was a Likert scale from 1 to 5, with anchors: 1 = Plotësisht kundër (Strongly disagree) and 5 = Plotësisht dakord (Strongly agree)

36. When decisions affect people's feelings, I prefer that AI systems consider the emotional consequences before proposing solutions.
37. I would feel more at ease if a decision proposed by AI was accompanied by an explanation of how the AI considered human aspects (emotions, well-being).
38. For decisions that affect people, I always prefer a person's decision, even if AI suggests a more 'efficient' solution.
39. I would be willing to allow AI to make routine decisions, but not decisions that have a major emotional impact on people.