DIGITALEDUCATIONREVIEW

Analyzing teacher candidates' arguments on AI integration in education via different chatbots

Volkan Duran[1]
[1] Iğdır Üniversitesi, Turkey, volkan.duran8@gmail.com, https://orcid.org/0000-0001-5573-4990

ABSTRACT

The burgeoning role of Artificial Intelligence (AI) in education prompts crucial discussions regarding its implications for teaching and learning. This qualitative study probes the argumentative perspectives of 118 teacher candidates from Iğdır University on the integration of AI into educational practices. Employing Toulmin's (1958) model, we analyzed their arguments, which encompass claims, evidence, warrants, backings, rebuttals, and conclusions, to ascertain their stance on AI's pedagogical integration. Utilizing four distinct AI chatbots—GPT-4, Gemini AI, Claude 3 Haiku, and Mistral AI—the research deciphers thematic undercurrents within these dimensions. Moreover, a novel methodological contribution is made through 'negative space exploration', focusing on the unmentioned themes to identify latent biases and assumptions in the argumentation. The study's dual analytical approach, combining AI-driven theme identification and negative space exploration, resulted in an enriched understanding of the content. Key findings suggest a nuanced perception among participants: while AI chatbots are acknowledged for enhancing educational efficiency and enabling personalized learning, concerns regarding diminished human interaction, potential erosion of critical thinking skills, and ethical use persist. The analyses also highlight the need for a balanced AI implementation that supports, not supplants, traditional educational methods. This research contributes to the ongoing debate on effective AI integration in education and calls for responsible pedagogical adoption of AI technologies.

KEYWORDS: AI in education, argumentation analysis, AI impact, Toulmin method, chatbots, negative space exploration.

1 INTRODUCTION

The integration of Artificial Intelligence (AI) into education heralds a transformative shift in pedagogical methodologies and learning experiences. AI technologies offer unprecedented opportunities for personalized learning, automation of administrative tasks, and insights into student performance and engagement patterns. This evolution towards an AI-enhanced educational landscape necessitates a comprehensive understanding of its implications, challenges, and potential to redefine teaching and learning paradigms. AI's role in education extends from intelligent tutoring systems and personalized learning environments to data-driven insights for educators and policymakers. The ability of AI to adapt to individual learning styles and provide real-time feedback presents a paradigm shift in instructional methodologies (Baker & Smith, 2014; Hwang et. al, 2020; Fahimirad & Kotamjani, 2018). However, the integration of AI in education also raises critical considerations regarding equity, privacy, and the ethical use of data (Holmes, Bialik, & Fadel, 2019). Ensuring that AI technologies augment rather than replace human interaction in education is paramount for fostering an inclusive and supportive learning environment (Weller, 2021).

The advancement of Artificial Intelligence (AI) technology has led to significant developments in the field of chatbots, incorporating machine learning and natural language processing to enhance their functionalities. This evolution has opened new avenues for academic research, particularly in the education sector, where chatbots are envisioned to play the role of smart teaching assistants. There is a growing encouragement for educators to adopt AI-based chatbots in classroom activities, recognizing their potential to contribute positively to the educational experience (Bibauw et al., 2019; Følstad & Brandtzæg, 2017; Hwang & Chang, 202; Kim et al., 2022; Malik et al., 2021; Tamayo et al., 2020; Sandu & Gide, 2019; Smutny & Schreiberova, 2020). Yet, the integration of AI into academia is not without its issues and hurdles. A significant worry is that AI technologies might reinforce pre-existing biases and discrimination within research and educational settings. Furthermore, there's a concern over the possibility of AI systems being exploited or tampered with, leading to outcomes that are either biased or not dependable (Kooli, 2023).

According to Toulmin (1958), an argument consists of claims supported by justified reasons. Alternatively, an argument is defined as comprising claims and the evidence supporting those claims. In his work, Toulmin (1958) outlined the main structures of the argumentation process as data-claim-warrant and identified its sub-components as backing-qualifier-rebuttal. Educational and developmental psychologists have been paying more and more attention to "argumentation" during the past fifteen years as a technique that might promote critical thinking and conceptual comprehension since psychologists can assess students' reasoning from a beneficial standpoint when they utilize argumentation. Students' arguments, whether they are presented in a written essay, a discussion, or another format, can be evaluated in a number of ways, such as simple or complicated, deep or shallow, balanced or unbalanced, supported or speculative, etc. A rising corpus of research demonstrates the relationship between scientific learning and conceptual shift and the caliber of students' arguments (Nussbaum, 2011). In order to resolve different opinions through argumentation, one must adopt

an analytical stance that reveals all relevant components of the topic (Walton, 2006) Important aspects of these debates include: taking a look at the issue from different angles, stating the position clearly, backing up the claims with evidence, using proper argumentation structures, and utilizing argumentation schemes (Aldağ, 2005). This method highlights the significance of an all-encompassing and multi-faceted examination in comprehending and successfully navigating the intricacies of arguments ((Duschl & Osborne, 2002; Erduran, Simon & Osborne, 2004; Jiménez-Aleixandre ve Erduran, 2007; Jonassen & Kim, 2010; Lazarou et al., 2016).

There has been some discussion about using AI chatbots in education. While proponents commend ChatGPT for its contribution to education, particularly in creating adaptive and personalized learning environments (Qadir, 2022), concerns have been raised by some academics regarding ethical considerations associated with ChatGPT's use (Mhlanga, 2023). For instance, in their study of ChatGPT, Farrokhnia et al. (2023) looked at its pros and cons. The pros included things like easier access to information, more personalized learning, and less teaching workload. The cons included things like worries about academic dishonesty, problems with response quality evaluation, and the possibility of bias and discrimination. Recently the integration of artificial intelligence (AI) chatbots has been used also to improve the argumentation skills in various contexts (Guo et. al, 2022; Guo et. al, 2023; Wambsganss et. al. 2021).

Argumentation is a complicated cognitive process that occurs in social circumstances with the goal of building knowledge or solving issues through the active use of language. Looking at the generating function of ChatGPT in relation to the linguistic, structural, and dialogic requirements of high-quality arguing articles (Su et. al, 2023),

The application of various AI-driven chatbots for analyzing student argumentations aligns perfectly with the goals of educational and developmental psychologists. It not only enriches the analysis by introducing a range of computational perspectives but also mirrors the complexities and nuances inherent in human cognitive processes. By leveraging AI chatbots, researchers can dissect the layers of student argumentations, assessing them for logical coherence, depth of understanding, and the ability to engage with counterarguments. This method offers a multifaceted view of student reasoning, significantly contributing to our understanding of how argumentation impacts learning and conceptual development.

Furthermore, this methodology plays a crucial role in uncovering both the gaps in students' understanding, the biases inherent in their reasoning, and the biases present within the chatbots themselves. It also identifies potential intervention points that could enhance deeper conceptual understanding. The insights derived from these analyses are invaluable for crafting more effective educational strategies and tools. These tools can be specifically designed to meet the unique needs and challenges uncovered through the AI-assisted examination of student arguments. Additionally, this process contributes to the refinement and improvement of educational chatbots. By pinpointing where chatbots may misinterpret or inadequately analyze student arguments, developers can work on enhancing the chatbots' understanding and interaction capabilities, thereby making them more effective as educational tools. This dual focus not only advances our comprehension of how students learn and reason but also propels forward the capabilities of AI in educational settings, ensuring that technology and pedagogy evolve in concert to better serve educational objectives.

2 METHODOLOGY

This is a qualitative research which is a case study based on content analysis by using four different chatbots. According to Yin (1999), the all-encompassing quality of a case study is its intense focus on a single event inside its real-life environment. This is why a case study is so comprehensive. The present research takes a qualitative approach, more specifically utilizing an exploratory case study design as its methodology. According to Yin (2014), the purpose of this exploratory case study was to analyze supposed causal linkages that are too complex to be investigated using standard surveys or experiments. According to Eisenhardt (1989) and Eisenhardt and Graebner (2007), case research is a method that can be utilized anytime the researcher is looking for answers to "why" and "how" questions. Merriam (1998) defines the case as a bounded system that is embodied as a person or thing that requires an in-depth investigation. Case studies are vital because they use to the complexity of reality and the embedded relationships of a number of factors that comprise reality in a given context. Boundedness is an essential component of case studies because it refers to the complexity of reality.

Content analysis is a technique used to examine the content of many types of data, including visual and linguistic data. It allows for the categorization of occurrences or events to facilitate their analysis and interpretation (Harwood & Garry, 2003).

The population consists from 118 students from various departments at Iğdır University in Turkey during the 2023-2024 academic year. This research seeks to uncover how future educators perceive AI's role in enhancing teaching and learning processes and how their views can be analyzed via different chatbots. In our study, typical case sampling was employed to select the study population of 118 teacher candidates. The selection of 118 teacher candidates from Iğdır University is pivotal as it encompasses a wide spectrum of perspectives within the teaching community. This heterogeneous sample captures a broad range of viewpoints, ensuring comprehensive insights into future educators' preparedness for AI integration in education.

The goal of typical case sampling is to "describe and show what is typical for people who are not familiar with the setting." People who are "key sources" or who use "statistical data" to find "average-like" cases choose typical cases. It is important to "try to get broad agreement about which cases are typical–and what criteria are being used to define typicality" when using typical case sampling (Suri, 2011). The "typical" is defined as representing a broad and diverse range of perspectives that future teacher candidates might have in there since the study aims to understand the general attitudes and preparedness of future teacher candidates towards AI integration in education, including candidates from different departments can provide a richer, more nuanced understanding. This diversity can highlight common themes and differences across various academic backgrounds, contributing to a more holistic view. Typical case sampling does not necessarily require homogeneity. The goal is to describe and illustrate what is average or common within the broader population. Including a range of departments can reflect the actual diversity within the population of teacher candidates, thus providing a realistic depiction of typical attitudes and preparedness levels. By including candidates from various academic branches who have completed pedagogy

courses, the study aims to represent the heterogeneous nature of the future teaching workforce. This diversity is crucial for capturing the typical case within this population, as it reflects the range of backgrounds and experiences that future educators bring to their roles. onsequently, this method enhances the validity and generalizability of the findings, ensuring they are applicable to a wide array of educational settings and contexts. Hence,

2.1 Background

Typical case sampling is instrumental in acquainting those not well-versed with the setting. For this research, this approach provided insights into the general attitudes and preparedness of prospective teacher candidates regarding AI integration in education. Candidates were selected from diverse branches who have completed pedagogy courses, particularly the principles and methods of instruction. This selection is pivotal as it encompasses a wide spectrum of perspectives within the teaching community, ensuring a heterogeneous sample that captures a broad range of viewpoints.

2.2 Purpose of Selection

The primary objective of selecting these teacher candidates is to assess their readiness for AI integration in education. Their insights are invaluable for understanding potential challenges and developing strategies for effective AI adoption in educational settings. The primary objective of typical case sampling is to depict what is standard within a given population. This study aims to capture the typical perspectives and arguments of future teacher candidates concerning the integration of AI in education. By selecting candidates from diverse academic branches who have undertaken pedagogy courses, it is ensured that the sample reflects the average experiences and viewpoints within this group. The main purpose of selecting these future educators is to assess their readiness for AI integration in education. Their insights are invaluable for gauging their preparedness and the potential challenges they might face. Additionally, since these candidates have completed pedagogy courses rather than solely coming from education faculties, their perspectives offer unique insights distinct from those of candidates from traditional education programs. This distinction provides a foundation for subsequent studies and improvements in teacher training programs.

2.3 Units of Analysis

The units of analysis include the views of individual teacher candidates, their academic backgrounds, and their participation in pedagogy courses. This approach provides a representative sample of future educators' attitudes and preparedness for integrating AI into education. The selection of these candidates, based on typical case sampling, ensures that they represent the average experiences and perspectives of future educators. Here, the specific criteria defining the units of analysis are elaborated:

2.3.1. Individual Teacher Candidates' Views

The views of each teacher candidate are considered as a unit of analysis. These views were selected because they provide a representative sample of future educators' attitudes and preparedness for integrating AI into education.

2.3.2. Academic Branches

The teacher candidates come from diverse academic disciplines (.Mathematics, Geography, English Language And Literature, Sports Sciences, Engineering, Music, Painting and Arts, History, Handicrafts, Public Relations) This diversity ensures that the sample captures a wide range of perspectives, reflecting the typical experiences of candidates from various educational backgrounds.

2.3.3. Participation in Pedagogy Courses

The selected teacher candidates have all completed pedagogy courses, specifically the principles and methods of instruction. This common educational experience is crucial as it provides a standardized basis for understanding their perspectives on AI integration in education.

2.3.4. Future Roles as Educators

The candidates are preparing for future roles as educators. This criterion ensures that their insights and attitudes towards AI are relevant and applicable to the practicalities of teaching and educational practices.

2.4 Rationale for Selecting These Units of Analysis

By defining the units of analysis as the views of individual teacher candidates with these specific characteristics, the following aims are pursued:

2.4.1. Capture Typical Experiences

Ensure that the views and perspectives analyzed are typical of those expected to be encountered in future educators. This helps in understanding the general trends and commonalities within this population.

2.4.2. Provide Comprehensive Insights

Gather detailed and nuanced insights into how future educators from various academic backgrounds and with pedagogical training view the integration of AI into education. This is critical for developing a holistic understanding of the readiness and potential challenges in this area.

2.4.3. Inform Educational Strategies

Use the findings from these typical cases to inform educational strategies, policies, and support mechanisms that can be implemented to better prepare future educators for the integration of AI in their teaching practices.

Utilizing the Toulmin (1958) method, participants constructed arguments in a semi-structured interview form that include claim, evidence, warrants, backings, rebuttal, and conclusion, assessing their stance towards AI in education by their argumentation including one Likert-type item measuring their attitude toward AI in education ranging from -10 to 1.

The content analysis was employed by both AI tools (GPT-4, Claude-AI, Claude 3 Haiku, Mistral AI) to identify themes, categories, and codes within the dimensions of claim, evidence, warrant, backing, rebuttal, and conclusion. The selected AI chatbots (GPT-4, Gemini-AI, Claude 3 Haiku, and Mistral AI) were chosen for their advanced language processing capabilities, versatility, and efficiency in handling large volumes of text. These chatbots offer unique strengths that contribute to a comprehensive qualitative analysis. GPT-4, or Generative Pre-trained Transformer 4, is the fourth iteration of OpenAI's large language model series. It is part of the GPT family, which consists of powerful neural network-based models designed to understand and generate human-like text. Gemini AI is Google's latest, most powerful large language model, capable of generating text, code, and even understanding images. It's a versatile tool that can help with writing, coding, problem-solving, and more. Built on a cutting-edge architecture

and trained on a vast dataset, Gemini promises to be a game-changer in the world of artificial intelligence, impacting everything from creative expression to scientific discovery. The Claude 3 model is another large language model. The model is capable of processing 21 thousand tokens per second for commands under 32 thousand tokens. This corresponds to approximately 30 pages. Mistral AI chat refers to a conversational AI system developed by Mistral AI. Mistral AI chat systems could be used in various applications, such as virtual assistants, customer support, content generation, and more. They are built using advanced machine learning techniques, particularly in the field of natural language processing. By selecting GPT-4, Gemini-AI, Claude 3 Haiku, and Mistral AI, it is aimed to harness the unique strengths and capabilities of each chatbot to conduct a comprehensive and multifaceted qualitative analysis. These chatbots were chosen because:

- They offer advanced language understanding and processing capabilities.
- They bring versatility in handling various types of data and perspectives.
- They provide efficiency in processing large volumes of text.
- They contribute to a detailed, nuanced, and thorough analysis of the teacher candidates' arguments.

This combination ensures that our study benefits from a robust and diverse analytical framework, enabling us to capture a wide range of insights and provide a well-rounded understanding of AI integration in education from the perspectives of future educators.

Firstly, chatbots are incorporated into the content analysis process, particularly by using four different chatbots to identify common categories, is an innovative approach to theme generation:

- Integration of Chatbots for Preliminary Analysis: Initially, the text is subjected to a preliminary analysis by four distinct chatbots. Each chatbot, with its unique processing capabilities and algorithms, reviews the text to identify key categories and themes present in the content.
- Extraction of Common Categories: Following the independent analyses by each chatbot, the next step involves collating the categories identified by all four AIs. The focus here is on finding overlapping categories or themes that are emphasized across the board, indicating a strong consensus among the different AI systems about the core elements of the text.
- Theme Creation Based on AI Consensus: The common categories identified by the chatbots serve as a foundation for theme development. These consensus categories are critically examined and synthesized into coherent themes that encapsulate the primary subjects or concepts present in the text, as interpreted by the AI systems with the help of human expert.

In the next step, the reverse analysis or negative space exploration was conducted, where the absence of information becomes the focal point in this study by also using chatbots. How this can be implemented and its potential benefits and challenges can be given as follows:

2.5 Implementation Steps

Analysis of omissions: following the identification of common themes by the chatbots, negative space exploration shifts the focus to what has not been mentioned in the text. this involves a deliberate search for gaps, omissions, and the absence of expected themes or categories that, by their absence, could offer insights into biases, assumptions, or underlying values within the text.

Comparative analysis: utilizing the themes and categories identified by the chatbots as a reference framework for identifying significant omissions. this step involves comparing the explicit content (as highlighted by the ai) against a broader expectation of what the content might logically include, based on the context, genre, or subject matter expertise.

Interpretation of negative space: the analysis of what is missing requires a deep understanding of the text's context, the potential biases of the chatbots, and an appreciation for the subtleties of communication. interpretations of negative space seek to uncover the unspoken assumptions, values, or priorities that shape the text.

Integration and synthesis: finally, integrate the findings from both the explicit themes identified by the chatbots and the implicit insights gained from negative space exploration. this synthesis provides a holistic view of the text, highlighting both what is said and what is left unsaid, thereby offering a more nuanced and comprehensive understanding.

Benefits of such an approach:

- Uncovering Hidden Biases: This method could help uncover biases or gaps in the original text by highlighting what is omitted. It provides a different perspective on the content, offering insights that might not be apparent from a traditional content analysis.
- Improving Chatbot Development: This exercise can help improve chatbot capabilities by challenging them to perform more complex and abstract forms of analysis, pushing the boundaries of natural language understanding and generation.

Limitation of such an approach

- Subjectivity and Interpretation: Identifying themes that are or are not mentioned involves a high degree of subjectivity and interpretation. Different chatbots might identify different sets of missing themes based on their perspectives and biases.
- Limitations of AI Understanding: Current AI models, including chatbots, may have limitations in understanding the deeper, contextual nuances of texts, which could affect their ability to accurately identify relevant but unmentioned themes.
- Complexity of Implementation: This approach requires sophisticated natural language processing capabilities and a nuanced understanding of the text's subject matter, making it challenging to implement effectively.

Adding negative space exploration to the analysis process, after utilizing multiple chatbots for theme identification, offers a sophisticated and comprehensive approach to content analysis. By examining both what is present and what is absent, one can achieve a deeper, more nuanced understanding of the text, uncovering underlying biases, assumptions, and the broader cultural or contextual implications. This method exemplifies a sophisticated blend of technology pushing the boundaries of traditional content analysis.

2.6 Validity and Reliability

In order to check the validity and reliability of the study between the student answers and the analysis of the chatbots for their answers, the answers of the students to one likert-type item measuring their attitude toward AI in education ranging from -10 to 10 are compared the ratings of the chatbots regarding their attitudes in their argumentation were examined in terms of correlation and Cronbach's Alpha values.

The table below presents Spearman's rho correlation coefficients among several variables related to the attitudes towards the use of artificial intelligence (AI) chatbots in education and the evaluations by four different AI chatbots: GPT-4, Gemini, Mistral, and Claude 3 Haiku. The coefficients aim to measure the strength and direction of association between these variables. The correlations between the claim that "using artificial intelligence chatbots in education increases the efficiency of education" and the evaluations by the four chatbots (GPT-4: 0.508, Gemini: 0.647, Mistral: 0.614, Claude 3 Haiku: 0.547) are all positive and statistically significant ($p < 0.01$). This indicates that respondents with a positive approach to the claim also tend to receive positive evaluations from the AI chatbots concerning this claim. Among the chatbots, Gemini shows the strongest correlation with the claim (0.647), suggesting that Gemini's evaluations are most closely associated with positive attitudes towards the efficiency of AI chatbots in education. The inter-chatbot evaluations also exhibit strong and significant positive correlations, particularly between Gemini and Mistral (0.807) and between Mistral and Claude 3 Haiku (0.791). This suggests a high level of agreement among these chatbots in evaluating attitudes towards AI in education. The particularly high correlations among the chatbots' evaluations (especially between Gemini and Mistral) highlight a potential for these AI tools to understand and align with human attitudes towards technology in education. This could be indicative of their sophisticated natural language processing and sentiment analysis capabilities. The analysis demonstrates a broad and significant positive reception towards the use of AI chatbots in education, both among respondents and as evaluated by the chatbots themselves. The strong correlations, especially between Gemini and Mistral, not only underscore the potential of AI chatbots in enriching educational experiences but also highlight the advanced capabilities of these chatbots in accurately capturing and reflecting human attitudes towards technological integration in education. This insight is valuable for educators, policymakers, and AI developers aiming to leverage AI chatbots to augment educational efficiency and engagement.

Correlations							
			Attitudes of the studetns	gpt4attitude	geminiattitude	mistralattitude	claude3haiku
Spearman's rho	Attitudes of the studetns	Correlation Coefficient	1,000	,508**	,647**	,614**	,547**
		Sig. (2-tailed)	.	,000	,000	,000	,000
		N		118	118	118	118
	gpt4attitude	Correlation Coefficient		1,000	,661**	,711**	,528**
		Sig. (2-tailed)			,000	,000	,000
		N			118	118	118
	geminiattitude	Correlation Coefficient			1,000	,807**	,769**
		Sig. (2-tailed)			.	,000	,000
		N				118	118
	mistralattitude	Correlation Coefficient				1,000	,791**
		Sig. (2-tailed)				.	,000
		N					118
**. Correlation is significant at the 0.01 level (2-tailed).							

Table 1. The correlation between students own attitutdes and evaluation of chatbots regarding their attitutdes

When we look at the Cronbach's Alpha that is a measure used to assess the internal consistency or reliability of a set of items or variables that are intended to measure the same underlying construct, I found the Cronbach's Alpha of 0.871 signifying that the set of five items exhibits high internal consistency. This level of reliability is considered excellent, as it exceeds the commonly accepted threshold of 0.7 for acceptable reliability. Values above 0.8 are often regarded as indicating good to excellent reliability, suggesting that the items measure the same underlying construct or concept effectively.

I outline comprehensive measures to ensure data privacy and ethical treatment, including anonymizing participant data and obtaining informed consent, especially regarding the use of AI in data analysis.

3 RESULTS

3.1 Descriptive Results

Based on the frequency table 2 provided to present data on responses to the claim "Using artificial intelligence chatbots in education increases the efficiency of education." the data suggests a slightly positive approach overall towards the claim, with the largest group (28.8%) strongly agreeing that using AI chatbots increases educational efficiency. However, there is also a significant proportion (39%) that disagreed with the claim. To summarize, the responses are somewhat divided, but there is a slightly more positive than negative approach to the claim based on this data sample.

Using artificial intelligence chatbots in education increases the efficiency of education. Do you have a positive or negative approach to this claim?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	20	16,9	16,9	16,9
	2	26	22,0	22,0	39,0

	3	22	18,6	18,6	57,6
	4	16	13,6	13,6	71,2
	5	34	28,8	28,8	100,0
	Total	118	100,0	100,0	

Table 2. Descriptive Values of the Attitude

Based on the new frequency table 3 provided, it - present data on the respondents' level of confidence in having a positive or negative approach to the claim "Using artificial intelligence chatbots in education increases the efficiency of education." The data shows that the largest group (30.5%) was extremely confident in their positive or negative approach to the claim about using AI chatbots for increasing educational efficiency. Looking at the cumulative percentages, 50% of respondents were at most moderately confident, while the other 50% were very confident or extremely confident in their stance. So while the responses were somewhat divided in terms of confidence levels, there was a slightly higher proportion (49.5%) that was very or extremely confident in their approach to the claim, compared to 33.9% who were not at all or only slightly confident.

In summary, based on this data, respondents tended to be more on the confident side regarding their positive or negative stance towards the claim, with the largest group being extremely confident.

Using artificial intelligence chatbots in education increases the efficiency of education. How confident are you that you have a positive or negative approach to this claim?					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	17	14,4	14,4	14,4
	2	23	19,5	19,5	33,9
	3	19	16,1	16,1	50,0
	4	23	19,5	19,5	69,5
	5	36	30,5	30,5	100,0
	Total	118	100,0	100,0	

Table 3. Descriptive Values of the Confidency

3.2 The Content Analysis of General Arguments by Chatbots

The table 4 categorizes and thematically summarizes positive aspects of AI integration in education in the student responses based on the codes by using specific AI technologies (GPT-4, Gemini 1.5, Claude 3 Haiku, Mistral 7b) as references.

GPT 4	**Gemini 1.5.**	**Claude 3 Haiku**	**Mistral 7b**	**Themes**
Positive Aspects				
Enhanced Learning Efficiency Motivation and Engagement	Increased Efficiency and Time Saving Enhanced Engagement and Motivation	Provide instant feedback to help students correct mistakes.	Provide additional resources and support for learners. can offer instant feedback and correction to learners, helping them to learn more effectively	"Optimized Learning Through Technology-Enhanced Methods"
Accessibility and Convenience	Improved Accessibility and Support	Offer 24/7 support by answering student questions.	Being available 24/7, providing learners with constant access to support and resources.	"Enhanced Learning Through Around-the-Clock Accessibility and Support"
	Personalized Learning	Personalize learning content based on student needs.	Personalize learning experiences based on individual learner needs.	"Customized Educational Experiences through Personalized Learning"
	Reduced Human Error	Track student progress to give teachers feedback.	can help teachers monitor learner progress and identify areas of improvement.	"Enhancing Educational Quality through Precision Monitoring"
Educational Equity				

Table 4. The Categories and Themes about the Positive Aspects of AI İntegration İn Education

The table 4 presents four main themes that capture the positive aspects of AI integration in education:

- Optimized Learning Through Technology-Enhanced Methods: Highlights how AI can improve learning efficiency and engagement by offering instant feedback and additional resources, making learning more effective and personalized.
- Enhanced Learning Through Around-the-Clock Accessibility and Support: Emphasizes the advantage of AI in providing 24/7 support and resources, improving accessibility and convenience for learners.
- Customized Educational Experiences through Personalized Learning: Discusses the capacity of AI to tailor learning experiences to individual student needs, enhancing the relevance and effectiveness of education.
- Enhancing Educational Quality through Precision Monitoring: Points out the benefit of AI in tracking student progress accurately, which helps in giving constructive feedback and identifying areas for improvement, thereby reducing human error.

The table 5 categorizes and thematically summarizes negative aspects of AI integration in education in the student responses based on the codes by using specific AI technologies (GPT-4, Gemini 1.5, Claude 3 Haiku, Mistral 7b) as references

Negative Aspects				
Diminished Human Interaction			May reduce the amount of human interaction and personal connection between learners and teachers.	
Impact on Critical Thinking and Creativity	Promotes Laziness and Reduces Effort Diminishes Creativity and Originality Superficial Learning and Lack	Can make students passive and reduce critical thinking.	Over-reliance on chatbots for learning could limit learners' creativity and critical thinking skills.	"The Double-Edged Sword of Technology in Learning: Impediments to Deep Engagement and Creative Thinking"

	of Retention			
Plagiarism and Academic Integrity Privacy and Ethical Concerns	Ethical Concerns and Potential for Cheating	May encourage cheating and plagiarism.		"Navigating the Ethical Pitfalls of Technology in Education: Integrity, Privacy, and Accountability"
	Lack of Human Connection and Emotional Suppor	Lack emotional support and empathy that teachers provide	Not be able to understand or respond to learners' emotional needs or provide empathy.	"The Digital Dilemma: Bridging the Emotional Gap in Technology-Enhanced Education"
Over-Reliance on Technology Questionable Reliability and Accuracy	Technical Issues and Accuracy Concerns	Can provide unreliable information leading to misinformation.	must be carefully considered and tested to ensure they are providing correct information.	"The Reliability Paradox in the Age of Digital Learning: Navigating the Pitfalls of Over-Reliance on Technology"
Loss of Traditional Educational Values				
Digital Divide				

Table.5. The Categories and Themes about the Negative Aspects of AI İntegration İn Education

The table 5 presents several main themes that capture the negative aspects of AI integration in education:

- Diminished Human Interaction: Notes the potential reduction in personal interaction and connection between learners and teachers due to AI's mediation.
- The Double-Edged Sword of Technology in Learning: Impediments to Deep Engagement and Creative Thinking: Warns about how an over-reliance on AI and technology can foster laziness, reduce effort, diminish creativity, lead to superficial learning, and limit critical thinking skills.
- Navigating the Ethical Pitfalls of Technology in Education: Integrity, Privacy, and Accountability: Raises concerns about academic integrity (e.g., cheating and plagiarism) and privacy issues associated with AI use in education.
- The Digital Dilemma: Bridging the Emotional Gap in Technology-Enhanced Education: Discusses the challenge of replicating the emotional support and empathy provided by human teachers in AI-driven educational environments.
- The Reliability Paradox in the Age of Digital Learning: Navigating the Pitfalls of Over-Reliance on Technology: Highlights concerns about the reliability and accuracy of information provided by AI, emphasizing the need for careful consideration and testing.

3.3 The Negative Space Exploration of the Results of the Content Analysis of General Arguments by Chatbots

The outlined themes that students did not mention in their argumentation of AI in education delve into critical and often overlooked dimensions of the integration of artificial intelligence within educational contexts. These themes highlight the complexity and breadth of considerations that must be addressed to responsibly navigate the adoption and implementation of AI technologies in schools and learning environments. The implications of these unmentioned themes can be given as below:

- Ethical Considerations: The argumentations does not deeply explore the ethical implications of AI in education, such as data privacy, the consent of minors, and how personal student data is used and protected.
- Long-term Impact on Employment: There's no discussion on how reliance on AI and technology in education might prepare or fail to prepare students for future job markets, especially in terms of automation and the need for digital literacy.
- Cultural and Social Implications: The impact of AI on cultural diversity in education, including whether AI can offer culturally responsive teaching or if it might lead to a homogenization of educational content.
- Accessibility for Students with Disabilities: The potential for AI to either enhance or hinder accessibility in education for students with disabilities, such as through adaptive learning technologies or barriers created by poorly designed interfaces.
- Cost Implications and Economic Barriers: The financial costs associated with implementing AI technologies in schools and whether this creates or exacerbates inequalities between different educational institutions or regions.
- Teacher and Educator Roles: A detailed examination of how the role of teachers and educators might evolve with the integration of AI, including potential shifts in responsibilities, skills, and job security.
- Psychological Impact on Students: The psychological effects of interacting with AI and technology in educational settings, including issues of student motivation, self-esteem, and the social dynamics of learning environments.

Each of these unmentioned themes could provide a deeper understanding of the complex relationship between AI technologies and educational outcomes, offering a more nuanced perspective on their integration into learning environments. Addressing these themes in discussions about AI in education is essential for developing a holistic understanding of the technology's potential benefits and challenges. It highlights the need for a multifaceted approach that considers ethical, social, cultural, economic, and psychological factors to ensure the responsible and effective integration of AI in learning environments.

The analysis provided by Gemini 1.5 brings to light a comprehensive set of themes concerning the integration of AI in education, which were not explicitly discussed in the original content of the argumentations of the students. These themes enrich the conversation by introducing critical dimensions related to ethical considerations, the impact on learning skills and styles, issues of accessibility and equity, and the long-term societal effects. The implications of these unmentioned themes can be given as below:

Ethical considerations

- Bias and discrimination: AI systems can perpetuate existing biases and discrimination if trained on biased data. This can lead to unfair outcomes for certain groups of students.

- Privacy concerns: AI systems collect large amounts of data on students, raising concerns about privacy and data security.
- The role of human teachers: The increasing use of AI in education raises questions about the future role of human teachers and the importance of human interaction in learning.

Impact on specific skills and learning styles:

- Creativity and critical thinking: While AI can help with rote learning and information retrieval, it is unclear how it affects the development of creativity and critical thinking skills.
- Collaboration and communication: The argumentationsmention the potential decrease in social interaction, but it could further explore how AI affects collaboration and communication skills.
- Adaptability and lifelong learning: The argumentations mention AI's ability to adapt to changing needs, but it could also discuss how AI can help students develop adaptability and lifelong learning skills.

Accessibility and equity:

- Digital divide: The argumentations briefly mention the issue of access to technology, but it could further explore how AI can exacerbate or help bridge the digital divide in education.
- Accommodating diverse learning needs: While the argumentations mention personalized learning, it could further discuss how AI can be used to accommodate diverse learning needs and disabilities.

Long-term impact on society:

- Changes in the job market: The argumentations mention the potential for AI to cause job losses, but it could also discuss how AI will change the skills needed in the future workforce and how education can prepare students for these changes.
- The impact on human relationships and social interaction: The argumentations mention the potential decrease in social interaction, but it could further explore the broader societal impact of increased reliance on AI in education and other areas of life.

These themes from Gemini 1.5's analysis present a holistic view of the multifaceted impact of AI in education, emphasizing the importance of ethical considerations, the development of a comprehensive skill set in students, ensuring accessibility and equity, and preparing society for long-term changes. Addressing these themes is crucial for harnessing AI's potential to enhance educational outcomes in a manner that is fair, inclusive, and sustainable.

Claude 3's insights contribute a valuable perspective to the discussion on the integration of AI in education, pinpointing critical areas that were not extensively covered in the initial analysis of the argumentations. The themes highlighted emphasize the ethical considerations, the effects on educational equity and accessibility, the role of teacher training, the impact on student learning competencies, the benefits for distance education, and the implications for student-teacher relationships. Here's a deeper look into each of these themes:

- Ethical aspects of using AI in education: Concerns about protecting students' personal data, AI producing biased or erroneous results, etc.
- Impact of AI in education on equity and accessibility: Ensuring all students can access AI-supported education, how socioeconomic differences affect access to educational opportunities.
- The need to train and support teachers in using AI: The necessity of providing teachers with the necessary support to effectively utilize AI technologies.
- How AI affects the development of students' creativity, critical thinking, and problem-solving skills: The impact of AI on the growth of these important competencies.
- The role and contributions of AI in distance education and online learning environments: The opportunities AI provides during the increased need for remote education during the pandemic.
- The impact of AI on student-teacher interaction and relationship: How AI affects the communication and bond between students and teachers.

Claude 3's analysis enriches the dialogue on AI in education by illuminating these essential themes, advocating for a balanced and thoughtful approach to technology integration. It underscores the importance of ethical considerations, equitable access, teacher empowerment, the development of student skills, the advantages for remote learning, and the preservation of meaningful student-teacher relationships. Addressing these themes is critical for realizing the full potential of AI in enhancing educational experiences and outcomes.

Mistral 7b's analysis brings to light several pivotal themes that were not addressed in the argumentations of the students regarding the usage of AI chatbots in education, highlighting the complexities and broader implications of deploying AI technologies in learning environments. These themes underscore the multifaceted nature of AI integration, touching on ethical concerns, the enhancement of social-emotional learning, accessibility issues, financial implications, the necessity of teacher training, long-term impacts, the need for regulatory measures, and the importance of accommodating cultural and linguistic diversity:

- Ethical concerns: The argumentations don't discuss ethical concerns related to the use of AI chatbots in education, such as privacy, data security, and potential biases in the algorithms used by the chatbots.
- Social-emotional learning: The argumentations don't mention the role of AI chatbots in supporting social-emotional learning, such as developing empathy, communication skills, and emotional intelligence.
- Accessibility: The argumentations don't discuss the potential of AI chatbots to make education more accessible to students with disabilities or those who live in remote areas.
- Cost: The argumentations don't mention the cost implications of using AI chatbots in education, such as the cost of developing and implementing the technology, and the potential impact on education budgets.
- Teacher training: The argumentations don't discuss the need for teacher training to effectively integrate AI chatbots into the

classroom and ensure that they are used in a pedagogically sound way.

- Long-term impact: The argumentations don't discuss the potential long-term impact of using AI chatbots in education on students' learning outcomes, career prospects, and overall development.
- Regulation and standardization: The argumentations don't mention the need for regulation and standardization of AI chatbots used in education to ensure quality, safety, and effectiveness.
- Cultural and linguistic diversity: The argumentations don't discuss the potential challenges and opportunities of using AI chatbots in education to support students from diverse cultural and linguistic backgrounds.

Mistral 7b's contribution to the analysis adds depth to the understanding of AI chatbots in education, suggesting areas for further exploration and development. These unmentioned themes underscore the need for a comprehensive approach to AI integration that considers ethical, social, financial, and educational factors to fully realize the benefits of AI for all students.

While there is some overlap in the themes identified by different chatbots, each one brings a unique perspective and highlights specific areas that deserve further exploration. The themes cover a wide range of issues, including ethical concerns, accessibility and equity, impact on skills development, cost implications, teacher training, long-term societal impact, and the need for regulation and standardization.

This analysis underscores the importance of considering the potential unmentioned or overlooked aspects when evaluating the integration of AI in education. By identifying these gaps, researchers, educators, and policymakers can engage in more comprehensive and nuanced discussions, addressing potential blind spots and ensuring a holistic understanding of the implications and considerations surrounding the use of AI in educational settings.

3.4 The Comparison of the analysis of four chatbots regarding the components of the students' argumentations

When comparing the analyses provided by four chatbots (GPT-4, Gemini 1.5, Claude 3, and Mistral 7b) regarding the components of students' argumentations—claim, evidence, warrants, backings, rebuttal, and conclusion—we delve into a nuanced examination of how each AI system potentially interprets and emphasizes different aspects of argumentative reasoning in an educational context. This comparison aims to shed light on the capabilities and focuses of each chatbot in analyzing argumentative structures, a fundamental skill in critical thinking and academic discourse.

3.4.1. The Comparison of the analysis of four chatbots regarding the claims of the students in their argumentations

Comparing the analyses of the four chatbots (GPT-4, Gemini 1.5, Claude 3 Haiku, and Mistral) on students' quality claims in their argumentation regarding the use of AI chatbots in education reveals distinct approaches and emphases in evaluating the argumentative reasoning process. Each analysis brings a unique perspective to the evaluation of students' argumentative skills, depth of understanding, critical thinking, and engagement with the topic.

GPT-4's Analysis:
GPT-4 offers a comprehensive evaluation, focusing on depth of understanding, evidence and reasoning, critical thinking, ethical and practical considerations, constructive alternatives, and varied perspectives. It commends students for their nuanced understanding and critical thinking but suggests areas for improvement, such as the need for more evidence-based claims, counterargument integration, and depth of reflection. GPT-4's analysis is thorough, providing a balanced view of the students' argumentative skills and their understanding of the complex implications of AI in education.

Gemini 1.5's Analysis:
Gemini 1.5 highlights specific arguments for and against the efficiency of AI chatbots in education, detailing the potential benefits and drawbacks identified by students. It emphasizes accessibility, personalized learning, and efficiency as key benefits, while also noting concerns about passivity, superficial learning, and academic integrity. Gemini 1.5's analysis is structured around the binary evaluation of efficiency, offering a detailed exploration of the arguments without delving deeply into the quality of argumentation or critical thinking processes.

Claude 3 Haiku's Analysis:
Claude 3 Haiku critiques the students' argumentation skills more directly, noting a general lack of skills in developing arguments, presenting evidence, and evaluating counterarguments. It acknowledges the diversity of opinions on the efficiency of AI chatbots but suggests that students struggle with the argumentation process. Claude's analysis is critical of the students' argumentative abilities, emphasizing the need for improvement in skill development and the careful evaluation of AI chatbots' use in education.

Mistral's Analysis:
Mistral offers a balanced view, recognizing the potential benefits of chatbots in providing instant feedback and personalized learning experiences, while cautioning against overreliance on technology. Mistral's analysis focuses on the nuanced understanding of chatbots' benefits and limitations, emphasizing the importance of complementing chatbots with human interaction and traditional learning methods. It suggests a thoughtful consideration of chatbots' role in education, advocating for further research to ensure their effective and ethical use.

Comparative Insights:

- Depth and Nuance: GPT-4 and Mistral provide more in-depth analyses, focusing on the nuanced understanding and critical thinking demonstrated by students. Both highlight the complexity of the issue and the need for a balanced approach to integrating AI in education.
- Argumentation Skills: Claude 3 Haiku is unique in directly critiquing students' argumentation skills, emphasizing the need for development in this area.

- Efficiency Focus: Gemini 1.5 uniquely structures its analysis around the concept of efficiency, providing a clear breakdown of arguments for and against the use of AI chatbots.
- Critical Thinking and Ethical Considerations: GPT-4 places significant emphasis on ethical and practical considerations, as well as the need for critical thinking and constructive alternatives, offering a more comprehensive evaluation of students' engagement with the topic.

In conclusion, each chatbot brings a distinct perspective to the evaluation, ranging from detailed analyses of argumentation skills and critical thinking to focused discussions on the efficiency and practical implications of AI chatbots in education. This comparison underscores the importance of a multifaceted approach to evaluating argumentative reasoning and the diverse aspects of integrating AI technology in educational contexts.

3.4.2. The Comparison of the analysis of four chatbots regarding the evidence of the students in their argumentations

The four analyses offer a spectrum of perspectives on evaluating students' argumentation skills, especially in the context of AI tools in education. Here's a comparative overview:

GPT-4's Analysis:
GPT-4 Analysis focuses on the depth and breadth of the arguments presented, including both supporting and critical views on AI in education. It emphasizes the importance of evidence diversity, critical thinking, and clarity in presentation, suggesting improvements in evidence specificity and argument clarity.

Gemini 1.5 Analysis:

Gemini 1.5's Analysis highlights the students' ability to use evidence-based arguments and address counterarguments. However, it points out weaknesses such as overgeneralization and logical fallacies. Recommendations for improvement include research on credible sources and teaching about logical fallacies.

Claude 3 Haiku's Analysis:

Claude 3 Haiku's Analysis seems to provide a balanced view of the positive and negative aspects of AI chatbots in education, emphasizing practical benefits and potential drawbacks. It suggests that while chatbots offer significant advantages in personalized learning and efficiency, they might also lead to reduced critical thinking and dependence on technology.

Mistral 7b Instruct's Analysis:

Mistral 7b Instruct's Analysis appears to focus more on the structure and substantiation of the arguments. It acknowledges the student's efforts to highlight AI benefits while noting the need for more structured argumentation and evidence to support claims. The analysis calls for a balanced view on the role of human interaction in education alongside AI tools.

Comparative Insights:
- Evidence and Structure: GPT-4 and Mistral 7b Instruct both emphasize the importance of well-structured arguments and the use of specific evidence. Gemini 1.5 also stresses the importance of credible evidence but places more focus on avoiding logical fallacies.
- Critical Thinking: Claude 3 Haiku and Gemini 1.5 highlight the importance of critical thinking, with Claude 3 pointing out the dual nature of AI tools' impact on education. Both suggest weighing both benefits and downsides.
- Recommendations for Improvement: Each analysis offers unique recommendations, from Gemini 1.5's advice on teaching logical fallacies to Mistral 7b's emphasis on argument structure. Claude 3 suggests focusing on the nuanced application of AI tools, considering both their potential and limitations.
- Overall Evaluation: While each analysis brings different criteria and focal points, they collectively underline the importance of critical thinking, evidence-based argumentation, and the nuanced understanding of AI tools in education. They advocate for a balanced approach to integrating technology in educational settings, ensuring that students not only embrace technological advantages but also remain aware of potential pitfalls.

This comparative analysis underscores the multifaceted nature of evaluating argumentation skills, highlighting the need for comprehensive criteria that include evidence quality, argument structure, critical thinking, and the ability to present a balanced view on contentious issues such as the integration of AI in education.

3.4.3. The Comparison of the analysis of four chatbots regarding the warrants of the students in their argumentations

The analyses of the four chatbots on students' argumentation skills regarding AI chatbots in education and their warrants (logical connections) show a nuanced understanding of both the potential benefits and drawbacks of integrating AI chatbots into educational settings. Let's compare these analyses:

GPT-4 Analysis:
Focuses on a broad range of warrants: This includes logical connections, the role of technology in education, the benefits of instant feedback and accessibility, concerns about reducing critical thinking and creativity, and the potential for personalized learning.

Considers ethical implications: It highlights concerns about academic honesty and the potential negative impacts on students' learning processes.

Emphasizes the need for balance and critical evaluation: The analysis suggests a nuanced approach to integrating AI chatbots in education, acknowledging both potential benefits and drawbacks.

Gemini 1.5 Analysis:
Outlines specific arguments for and against AI chatbots: It directly contrasts the potential efficiency and personalization benefits with concerns about reduced critical thinking, creativity, and dependence on technology.

Cites the need for evidence-based decision-making: Gemini 1.5 calls for further research and evidence to determine the true impact of AI chatbots on education, emphasizing a pragmatic approach.

Addresses ethical and reliability concerns: Similar to GPT-4, it highlights the importance of addressing ethical concerns and the accuracy of information provided by chatbots.

Claude 3 Haiku's Analysis:
Employs a structured approach to warrants: It identifies specific types of warrants used by students, including statistical data, expert opinions, examples, cause-and-effect relationships, and personal experiences.

Highlights both positive and negative views: Claude 3 Haiku presents a balanced view of AI chatbots' role in education, acknowledging their potential to enhance learning while also recognizing the risks of dependency and reduced critical thinking skills.

Suggests a nuanced perspective: By examining the types of evidence and logical connections used by students, this analysis fosters a deeper understanding of the complexities involved in the debate over AI chatbots in education.

Mistral 7b Instruct's Analysis:
Focuses on the logical structure of arguments: This analysis delves into the logical connections students make between their claims and the evidence, highlighting the importance of a well-reasoned argument.

Emphasizes the variety of evidence: It notes the use of statistical data, expert opinions, analogies, illustrations, and cause-and-effect relationships to support claims about AI chatbots in education.

Points out the consideration of counterarguments: Mistral 7b Instruct recognizes the students' ability to acknowledge and address potential criticisms of their positions, showing a mature approach to argumentation.

Comparative Insights:

- All four analyses recognize the complexity of integrating AI chatbots in education, citing both potential benefits and concerns. They emphasize the importance of logical connections between claims and evidence, ethical considerations, and the need for a balanced, evidence-based approach. However, they differ in their emphasis:
- GPT-4 and Gemini 1.5 offer a broader analysis of the ethical implications and the need for balance in evaluating the use of AI chatbots.
- Claude 3 Haiku and Mistral 7b Instruct focus more on the structure of the argumentation and the types of warrants used, offering a more detailed analysis of the logical underpinnings of students' arguments.
- Overall, these analyses together provide a comprehensive overview of the considerations involved in discussing AI chatbots in education, highlighting the importance of critical thinking, ethical considerations, and the careful evaluation of both the potential benefits and drawbacks of technological integration.

3.4.4. The Comparison of the analysis of four chatbots regarding the backings of the students in their argumentations

When comparing the analyses regarding the backings of students' argumentations across the four chatbots, we observe nuanced approaches to evaluating how students support their claims about the integration and impact of AI chatbots in education. Each analysis provides a distinct perspective, yet they collectively highlight the complexity and multi-faceted nature of the debate. Let's summarize and compare these analyses:

GPT-4 Analysis:

Depth and Nuance: GPT-4 offers a detailed analysis, categorizing the types of backings into seven distinct themes, ranging from statistical data and expert opinions to ethical considerations.

Comprehensive Coverage: It addresses both the technological advancements and the broader implications of AI in education, emphasizing a balanced view that considers potential benefits alongside ethical and societal concerns.

Gemini 1.5 Analysis:

Strengths and Weaknesses: Gemini focuses on the variability in the strength of backings, distinguishing between strong evidence (like research findings) and weaker forms (such as anecdotal evidence).

Critical Evaluation: This analysis suggests the importance of critically evaluating the evidence, highlighting the role of source, relevance, and evidence strength in determining the validity of students' claims.

Claude 3 Haiku's Analysis:

Diverse Evidence: Similar to GPT-4, Claude identifies a range of backings used by students, from statistical data to expert opinions and real-life examples.

Acknowledgment of Counterarguments: It uniquely points out the existence of counterarguments and the use of evidence to challenge the prevailing optimism about AI in education, such as concerns about reliance on chatbots reducing critical thinking skills.

Mistral 7B Instruct's Analysis:

Focus on Educational Effectiveness: Mistral delves into the educational effectiveness of AI and chatbots, noting how students use backings to argue for improvements in learning outcomes.

Consideration of Challenges: It also acknowledges challenges and limitations, such as technological access disparities and the potential impact on students' learning habits and critical thinking skills.

Comparative Insights:

- Varied Focus: While GPT-4 and Claude provide a broad overview of the types of backings and the ethical dimensions, Gemini emphasizes the critical evaluation of evidence strength, and Mistral zeroes in on educational effectiveness and challenges.
- Common Themes: Across all analyses, there is a consensus on the importance of diverse and reliable evidence in supporting claims, including statistical data, expert opinions, and real-world applications.
- Consideration of Challenges: Each analysis, in varying degrees, acknowledges the potential drawbacks and limitations of integrating AI chatbots into education, suggesting a nuanced understanding of the issue.
- In conclusion, these analyses collectively underscore the richness of student argumentations on AI chatbots in education. They highlight the importance of well-supported claims, the need for critical evaluation of evidence, and the necessity of considering both the potential benefits and challenges of AI integration in educational contexts. Together, they provide a comprehensive picture of how students back their arguments, emphasizing the depth and diversity of perspectives in this ongoing debate.

3.4.5. The Comparison of the analysis of four chatbots regarding the rebuttals of the students in their argumentations

The four analyses offer a diverse perspective on student rebuttals regarding the use of AI chatbots in education, each focusing on different aspects of the debate:

GPT-4 Analysis:

This analysis offers a comprehensive overview of the rebuttals, methodically categorizing them into concerns about human interaction and empathy, technical limitations, ethical and societal implications, impact on learning habits, accessibility issues, the need for human guidance, and potential for misuse. It emphasizes the critical perspective on AI chatbots, highlighting the need for careful consideration and human oversight.

Gemini 1.5 Analysis:

Gemini takes a more balanced approach by presenting both arguments for and against AI chatbots. It outlines the potential benefits such as increased efficiency and improved learning while also acknowledging counterarguments related to reduced human interaction, limited understanding and empathy, technical issues, and the potential for abuse. Gemini suggests a need for further research and responsible implementation.

Claude 3 Haiku's Analysis:

Focusing on the negative implications of AI chatbots, Claude 3 Haiku echoes concerns similar to those raised by GPT-4, such as the potential reduction in human interaction, lack of empathy, technical disruptions, misinformation risks, and the hindrance of creativity and critical thinking skills. This analysis also raises the issue of AI leading to laziness and a decline in social interaction among students.

Mistral 7B Instruct's Analysis:

This analysis combines specific rebuttals against the effectiveness of AI chatbots with a broader reflection on their impact. It mentions concerns about the lack of human touch, potential misinformation, encouragement of laziness, and various risks associated with overreliance on technology. It also hints at the need for a balanced and cautious approach towards integrating AI chatbots into education.

Comparative Insights:

- Depth of Analysis: GPT-4 offers the most thorough and organized examination of the rebuttals, providing a comprehensive understanding of the concerns related to AI chatbots in education.
- Balanced Perspective: Gemini 1.5 stands out for presenting a balanced view, acknowledging the benefits of AI chatbots while also considering the counterarguments, suggesting a nuanced approach to their implementation.
- Focus on Negative Implications: Both Claude 3 Haiku and Mistral 7B Instruct focus more on the potential negative impacts of AI chatbots, emphasizing concerns about the loss of human interaction and the risk of fostering dependency and laziness among students.
- Suggested Approach: While GPT-4 and Gemini 1.5 lean towards suggesting careful consideration and responsible implementation, Claude 3 Haiku and Mistral 7B Instruct call for a cautious approach, highlighting the importance of not overlooking the human elements of education in favor of technological convenience.
- Overall, the analyses collectively underscore the complexity of integrating AI chatbots into educational settings. They highlight the need for ongoing dialogue, ethical consideration, and empirical research to navigate the potential benefits and challenges of this technological integration, ensuring that it complements rather than undermines the human aspects of learning and teaching.

3.4.6. The Comparison of the analysis of four chatbots regarding the conclusions of the students in their argumentations

When comparing the analyses of four different chatbot responses regarding the conclusions drawn by students on AI chatbots in education, it becomes evident that each analysis presents a nuanced understanding of the potential impacts, benefits, and challenges posed by AI chatbots in educational settings. Here's a comparative overview:

Common Themes Across Analyses:

- Positive Aspects of AI Chatbots: All analyses recognize the potential benefits of AI chatbots in education, including increased efficiency, personalized learning, accessibility, and the capacity to provide immediate support and information.
- Concerns About Dependency and Loss of Critical Skills: There is a consensus across the analyses about the concern that reliance on chatbots may reduce students' motivation to engage in independent research, critical thinking, and problem-solving skills.
- Skepticism on Effectiveness and Ethical Implications: Each analysis raises questions about the reliability and depth of the information provided by chatbots. Concerns about ethical implications, such as data privacy and academic integrity, are also commonly expressed.
- The Necessity for Balanced and Responsible Implementation: A recurring recommendation is for a balanced approach to integrating chatbots, suggesting they should support, not replace, traditional educational methods and human interaction.
- Calls for Further Research: There's an acknowledgment of the need for empirical research to better understand the effects of chatbots on learning outcomes and the educational process.

Divergent Perspectives and Emphasis:

- Optimism for Technological Support: GPT-4's analysis particularly emphasizes optimism for the potential of AI to complement human teaching and enhance educational accessibility, reflecting a somewhat more positive stance on the technological support provided by chatbots.
- Detailed Concerns Over Human Element and Dependency: Claude 3 Haiku's analysis specifically highlights how chatbots may foster laziness and erode research and thinking skills, indicating a deeper focus on the potential negative impact of diminished human interaction and the development of critical skills.

- Mixed Views on Educational Outcomes: Gemini 1.5 and Mistral 7B Instruct's analyses provide a balanced view on both the potential improvements in learning efficiency and the risks of superficial engagement, highlighting the diversity of student opinions on educational outcomes.
- Specificity in Student Concerns and Recommendations: While all analyses discuss the importance of a balanced approach and further research, the level of detail and the specific concerns and recommendations vary. For example, Claude 3 Haiku and Mistral 7B Instruct tend to provide more specific insights into the potential misuse of chatbots and the importance of ethical considerations.

In conclusion, while there are common themes in the analyses concerning the integration of AI chatbots in education, the emphasis on certain aspects varies, reflecting a complex and multifaceted debate among students. These differences underscore the importance of considering a wide range of perspectives when evaluating the potential of AI chatbots in educational contexts, highlighting both the promise and the pitfalls of this technology.

4 DISCUSSION

This study underscores the efficacy of integrating chatbots into the content analysis process, particularly emphasizing the innovative approach of combining AI-driven preliminary theme identification with negative space exploration. This dual-methodology not only enables the identification of explicit themes within texts but also delves into the implicit insights by examining the absence of information, thereby offering a comprehensive understanding of the content under study. The employment of multiple chatbots for the initial theme generation ensures a broad and nuanced identification of categories, which, when coupled with human analysis for theme refinement and contextualization, leads to a sophisticated synthesis of AI efficiency and human insight.

Moreover, the unique application of negative space exploration adds a new dimension to content analysis, unveiling underlying biases, assumptions, and values that might not be immediately apparent, thereby enriching the analysis with deeper insights that traditional methods might overlook.

A significant part of this study's contribution is its demonstration of how chatbots can accurately reflect student attitudes towards AI in education, as evidenced by the Spearman's correlation coefficients. This empirical evidence showcases the potential of chatbots not only in identifying themes but also in accurately capturing and reflecting complex human sentiments, with Gemini 1.5 demonstrating a stronger correlation to actual student attitudes than GPT-4. This finding highlights the potential differences in AI capabilities and their application in nuanced data interpretation, marking a significant step forward in educational research methodologies.

The integration of artificial intelligence (AI) chatbots into education sparks a multifaceted debate that underscores the complexities of technological advancements in learning environments. This discussion synthesizes the descriptive results, content analysis, negative space exploration, and comparative analysis of students' argumentations derived from the engagement with various AI chatbot technologies such as GPT-4, Gemini 1.5, Claude 3 Haiku, and Mistral 7b. It reveals a landscape punctuated by optimism for enhanced learning efficiency and concerns over the erosion of critical educational elements.

The data suggests a landscape of educational innovation where AI chatbots are seen as a double-edged sword, offering potential improvements in educational efficiency and personalized learning, yet raising valid concerns about diminished human interaction and the erosion of critical thinking skills. This paradox reflects the ambivalence within the educational community about embracing AI technologies. While some embrace the convenience and adaptive potential of chatbots, others caution against a tech-centric approach that might overlook the nuanced demands of a comprehensive educational experience.

The positive aspects of AI chatbots as reflected in student responses—such as enhanced learning efficiency, around-the-clock support, and personalized learning experiences—align with the evolving demands of a 21st-century education system that values flexibility, accessibility, and individualized learning paths. These attributes, underscored by the analysis of AI chatbots themselves, suggest a shift in educational paradigms where traditional barriers are being challenged by the capabilities of AI technologies.

However, this technological optimism is balanced by concerns about the reliability and ethical use of AI chatbots. The risk of promoting superficial learning, fostering dependency, and compromising academic integrity presents a set of challenges that the education system must address. Moreover, the digital divide and issues of equitable access highlight the broader social implications of integrating AI into education. These concerns are amplified by the observations that critical student perspectives on such matters were often unvoiced or underrepresented in their argumentations, pointing to areas where deeper inquiry and more inclusive dialogue are necessary.

The nuanced approach of chatbots like Gemini 1.5, which revealed unmentioned themes including potential biases, privacy concerns, and the long-term impact on employment, aligns with current discussions about AI ethics and the future of work. Claude 3 Haiku's contributions further emphasize the need for balance, particularly in preserving the emotional and social aspects of learning that AI cannot fully replicate.

The comparison of chatbot analyses regarding the components of student argumentations—claims, evidence, warrants, backings, rebuttals, and conclusions—reveals both the diversity and commonality of student attitudes. While all chatbots recognize the potential benefits of AI chatbots in enhancing learning experiences, they offer different levels of critique regarding the potential downsides.

In drawing conclusions, it is evident that the integration of AI chatbots in education is not a simple matter of technological implementation; it involves a careful consideration of pedagogical, ethical, and social dimensions. The data illustrates that while students are generally confident in their attitudes towards AI chatbots, their level of confidence does not necessarily correlate with an in-depth understanding of the implications of such integration. This highlights an opportunity for educational institutions to foster deeper critical engagement with the potential impacts of AI technologies on educational practices and outcomes.

Ultimately, this study sheds light on the multifaceted nature of AI's role in education. It underscores the need for ongoing, critical, and inclusive conversations that engage various stakeholders—students, educators, technologists, and policymakers—in navigating the promising yet complex landscape of AI in education. The findings call for a prudent approach that harnesses the

strengths of AI chatbots while remaining cognizant of their limitations, ensuring that the educational journey remains both human-centric and future-ready.

5 CONCLUSIONS

The research findings highlight the innovative and effective integration of chatbots into the content analysis process, marking a significant advancement in research methodologies. This novel approach enhances thematic analysis by uncovering both explicit and implicit content, thereby facilitating a more nuanced and comprehensive understanding of texts. Furthermore, the study validates the application of chatbots in educational research, evidenced by their ability to accurately mirror student attitudes towards AI in education.

The introduction of negative space exploration as a technique in content analysis represents a groundbreaking advancement, offering a new perspective on text analysis that can reveal critical insights into underlying biases and assumptions. This method has broad implications, suggesting its applicability in various research fields where the analysis of unmentioned or overlooked aspects can provide profound insights.

However, the study also draws attention to the challenges and limitations associated with the integration of AI into content analysis, such as the nuances of AI understanding, the subjective nature of theme identification, and the complexity of effectively implementing these technologies. These challenges underscore the need for ongoing refinement of both AI technologies and analytical methodologies to ensure they enhance rather than replace human analytical capabilities.

The empirical evidence showcasing chatbots' ability to accurately reflect student attitudes towards AI in education demonstrates their potential not only in identifying themes but also in capturing complex human sentiments. This is exemplified by Gemini 1.5's stronger correlation to actual student attitudes compared to GPT-4 (author, year), highlighting the potential differences in AI capabilities and their application in nuanced data interpretation.

The study reveals a landscape where AI chatbots are perceived as a double-edged sword, offering potential improvements in educational efficiency and personalized learning, yet raising valid concerns about diminished human interaction and the erosion of critical thinking skills. This paradox reflects the ambivalence within the educational community about embracing AI technologies (author, year).

In conclusion, this study not only exemplifies the rich interplay between technology and human insight in the realm of research but also provides valuable, innovative methodologies that can be leveraged across disciplines. By exploring the capabilities of AI while critically engaging with its limitations, researchers can push the boundaries of content analysis, unlocking richer and more comprehensive insights into their subjects of study, and highlighting the pivotal role of AI in advancing research methodologies. The discourse surrounding the integration of AI chatbots in education is characterized by both enthusiasm and caution. It highlights the potential for AI to revolutionize education by making learning more efficient, personalized, and accessible. However, it also raises critical questions about the implications of such technological advancements for the human elements of education. Addressing these concerns through responsible implementation, ongoing research, and inclusive dialogue will be essential for harnessing the potential of AI chatbots in enhancing educational outcomes while safeguarding the integrity and humanity of the learning experience.

REFERENCES

Aldağ, H. (2005). *The effects of textual and graphical-textual argumentation software as cognitive tools on development of argumentation skills* [Unpublished doctoral dissertation]. Adana.

Baker, R., & Siemens, G. (2014). Educational data mining and learning analytics. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed., pp. 443-464). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.016

Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, 32(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508

Duschl, R., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, 38, 39-72. https://doi.org/10.1080/03057260208560187

Eisenhardt, K.M. (1989). Building theories from case study research. *Academy of Management Review*, 14 (4), 532-550. https://doi.org/10.2307/258557

Eisenhardt, K.M. & Graebner, M.E. (2007). Theory building from cases: Opportunities and challenges, *Academy of Management Journal*, 50 (1), 25-32

Erduran, S., Simon, S., & Osborne, J. (2004). TAPping into argumentation: Developments in the application of Toulmin's argument pattern for studying science discourse. *Science Education*, 88, 915-933. https://doi.org/10.1002/sce.20012

Fahimirad, M., & Kotamjani, S. S. (2018). A review on application of artificial intelligence in teaching and learning in educational contexts. *International Journal of Learning and Development*, 8(4), 106-118. https://doi.org/10.5296/ijld.v8i4.14057

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*. Advance online publication. https://doi.org/10.1080/14703297.2023.2195846

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, 24(4), 38–42. https://doi.org/10.1145/3085558

Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing*, 54, 100666. https://doi.org/10.1016/j.asw.2022.100666

Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education*, 203, 104862. https://doi.org/10.1016/j.compedu.2023.104862

Harwood, T. G., & Garry, T. (2003). An overview of content analysis. *The Marketing Review*, 3(4), 479-498. https://doi.org/10.1362/146934703771910080

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign. https://doi.org/10.58863/20.500.12424/4276068

Hwang, G., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artificial intelligence in education. *Computers & Education: Artificial Intelligence*, 1, [100001]. https://doi.org/10.1016/j.caeai.2020.100001

Jiménez-Aleixandre, M. P., & Erduran, S. (2007). Argumentation in science education: An overview. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentation in Science Education* (pp. 3-28). Springer Science + Business Media B.V. https://doi.org/10.1007/1-4020-6051-6_1

Kim, H., Yang, H., Shin, D., & Lee, J. H. (2022). Design principles and architecture of a second language learning chatbot. *Language Learning & Technology*, 26(1), 1-18. http://hdl.handle.net/10125/73463

Kooli, C. (2023). Chatbots in Education and Research: A Critical Examination of Ethical Implications and Solutions. *Sustainability*, 15(7), 5614. https://doi.org/10.3390/su15075614

Malik, R., Shrama, A., Trivedi, S., & Mishra, R. (2021). Adoption of chatbots for learning among university students: Role of perceived convenience and enhanced performance. *International Journal of Emerging Technologies in Learning (iJET)*, 16(18), 200-212. https://doi.org/10.3991/ijet.v16i18.24315

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. Jossey-Bass.

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.4354422

Nussbaum, E. M. (2011). Argumentation, Dialogue Theory, and Probability Modeling: Alternative Frameworks for Argumentation Research in Education. *Educational Psychologist*, 46(2), 84–106. https://doi.org/10.1080/00461520.2011.558816

Qadir, J. (2022). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. *TechRxiv*. https://doi.org/10.36227/techrxiv.21789434.v1

Sandu, N., & Gide, E. (2019, September). Adoption of AI-Chatbots to enhance student learning experience in higher education in India. In 2019 *18th International Conference on Information Technology Based Higher Education and Training (ITHET)* (pp. 1-5). IEEE. https://doi.org/10.1109/ITHET46829.2019.8937382

Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the Facebook Messenger. *Computers & Education*, 151. https://doi.org/10.1016/j.compedu.2020.103862

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. https://doi.org/10.1016/j.asw.2023.100752

Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative Research Journal*, 11(2), 63-75. https://doi.org/10.3316/QRJ1102063

Tamayo, P. R. V., Herrero, A. M., Martín, J., Navarro, C., & Tránchez, J. M. (2020). Design of a chatbot as a distance learning assistant. *Open Praxis*, 12(1), 145. https://doi.org/10.5944/openpraxis.12.1.1063

Toulmin, S. (1958). *The Uses of Argument*. Cambridge University Press.

Walton, D. (2006). Examination dialogue: An argumentation framework for critically questioning an expert opinion. *Journal of Pragmatics*, 38(5), 745-777. https://doi.org/10.1016/j.pragma.2005.01.016

Wambsganss, T., Guggisberg, S., & Söllner, M. (2021). ArgueBot: A conversational agent for adaptive argumentation feedback. *In Innovation Through Information Systems* (pp. 267-282). Springer. https://doi.org/10.1007/978-3-030-85868-4_18

Weller, M. (2021). *25 Years of Ed Tech*. Athabasca University Press.

Yin, R. K. (1999). Enhancing the quality of case studies in health services research. *Health Services Research*, 34 (5 Pt 2), 1209-1224. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1089060/

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Sage.

ANÀLISI DELS ARGUMENTS DELS CANDIDATS A DOCENTS SOBRE LA INTEGRACIÓ DE LA IA A L'EDUCACIÓ MITJANÇANT DIFERENTS CHATBOTS

El paper creixent de la Intel·ligència Artificial (IA) en l'educació planteja discussions crucials sobre les seves implicacions per a l'ensenyament i l'aprenentatge. Aquest estudi qualitatiu examina les perspectives argumentatives de 118 candidats a docents de la Universitat d'Iğdır respecte a la integració de la IA en les pràctiques educatives. Utilitzant el model de Toulmin (1958), analitzem els seus arguments, que abasten afirmacions, evidències, garanties, suports, refutacions i conclusions, per determinar la seva postura sobre la integració pedagògica de la IA. Emprant quatre xatbots de IA diferents —GPT-4, Gemini-AI, Claude 3 Haiku i Mistral AI—, la recerca desxifra les corrents temàtiques dins d'aquestes dimensions. A més, es realitza una nova contribució metodològica a través de l'"exploració de l'espai negatiu", centrant-se en els temes no esmentats per identificar biaixos i suposicions latents en l'argumentació. El doble enfocament analític de l'estudi, que combina la identificació de temes impulsada per la IA i l'exploració de l'espai negatiu, ha resultat en una comprensió enriquida del contingut. Els resultats claus suggereixen una percepció matissada entre els participants: tot i que es reconeix els xatbots de IA per millorar l'eficiència educativa i possibilitar l'aprenentatge personalitzat, persisteixen les preocupacions respecte a la disminució de la interacció humana, l'erosió potencial de les habilitats de pensament crític i l'ús ètic. Les anàlisis també destaquen la necessitat d'una implementació equilibrada de la IA que recolzi i no reemplaci els mètodes educatius tradicionals. Aquesta investigació contribueix al debat continu sobre la integració efectiva de la IA en l'educació i advoca per una adopció pedagògica responsable de les tecnologies de IA.

PARAULES CLAU: IA en educació, Anàlisi de l'argumentació, impacte de la IA, mètode de Toulmin, chatbot, exploració espacial negativa.

ANÁLISIS DE LOS ARGUMENTOS DE LOS CANDIDATOS A DOCENTES SOBRE LA INTEGRACIÓN DE LA IA EN LA EDUCACIÓN A TRAVÉS DE DIFERENTES CHATBOTS

El creciente papel de la Inteligencia Artificial (IA) en la educación suscita discusiones cruciales acerca de sus implicaciones para la enseñanza y el aprendizaje. Este estudio cualitativo examina las perspectivas argumentativas de 118 candidatos a docentes de la Universidad de Iğdır sobre la integración de la IA en las prácticas educativas. Utilizando el modelo de Toulmin (1958), analizamos sus argumentos, que abarcan afirmaciones, evidencias, garantías, respaldos, refutaciones y conclusiones, para determinar su postura sobre la integración pedagógica de la IA. Empleando cuatro chatbots de IA distintos —GPT-4, Gemini-AI, Claude 3 Haiku y Mistral AI—, la investigación descifra las corrientes temáticas dentro de estas dimensiones. Además, se realiza una novedosa contribución metodológica a través de la 'exploración del espacio negativo', enfocándose en los temas no mencionados para identificar sesgos y suposiciones latentes en la argumentación. El doble enfoque analítico del estudio, que combina la identificación de temas impulsada por IA y la exploración del espacio negativo, resultó en una comprensión enriquecida del contenido. Los hallazgos clave sugieren una percepción matizada entre los participantes: si bien se reconoce a los chatbots de IA por mejorar la eficiencia educativa y posibilitar el aprendizaje personalizado, persisten las preocupaciones con respecto a la disminución de la interacción humana, la posible erosión de las habilidades de pensamiento crítico y el uso ético. Los análisis también resaltan la necesidad de una implementación equilibrada de la IA que apoye y no reemplace los métodos educativos tradicionales. Esta investigación contribuye al debate continuo sobre la integración efectiva de la IA en la educación y aboga por una adopción pedagógica responsable de las tecnologías de IA.

PALABRAS CLAVE: IA en educación, Análisis deargumentación, impacto de la IA, método de Toulmin, chatbot, exploración espacial negativa