

# Human researcher vs. AI-supported qualitative data analysis: Hybrid prompt design for constant comparison analysis

Nurşin Akman <sup>1</sup>, Erman Uzun <sup>1</sup>, Pinar Arslan <sup>2</sup>

<sup>1</sup> Computer Education & Instructional Technology, Mersin University, Mersin, Türkiye; <sup>2</sup> Ministry of National Education, Niğde, Türkiye

## ABSTRACT

This study examines the potential integration of artificial intelligence (AI) tools, particularly ChatGPT, in qualitative data analysis educational research. The research tries to clarify AI-supported data analysis processes in qualitative data from teacher interviews that examine deficiencies in a "Fundamentals of Programming" course in vocational high schools. The methodology employs a hybrid (both heuristic and literature based) prompt engineering strategy, utilizing open, axial, and selective coding, to ensure a comprehensive analysis. The aim was to compare human and AI-supported analysis to evaluate the depth, efficiency, and reliability of AI tools in qualitative research. This study chose a comparative case study design because examining programming instruction with two different data analysis methods (human researcher and AI-supported) requires a comparative perspective. The findings indicate that hybrid prompt design for AI can significantly enhance the efficiency and accuracy of qualitative data analysis, providing deeper insights and more structured outputs. However, the study also highlights the importance of carefully designed prompts and human oversight to mitigate potential biases and errors inherent in AI-supported analysis. This research contributes to the growing field of AI in data analysis, offering a framework for future studies to leverage AI technologies for qualitative data analysis, thereby enhancing research quality and productivity.

## ARTICLE HISTORY

Received 27 November 2024

Accepted 20 February 2025

## KEYWORDS

AI-supported data analysis

Hybrid prompt design

Artificial intelligence

Qualitative data analysis

Constant comparison analysis

## CORRESPONDENCE

Erman Uzun



## Introduction

Technology plays a crucial role in enhancing teaching and learning processes. The advancement of artificial intelligence (AI) has facilitated the implementation of personalized teaching and assessment methods. AI-supported tools enable teachers and students to address their needs more efficiently, contributing to innovation in education (Göçmez, 2023). However, integrating these technologies into the educational environment requires technical knowledge and pedagogical understanding (Lima et al., 2023).

The "Fundamentals of Programming" course offered in vocational high schools aims to provide students with basic programming knowledge in information technologies. Effective teaching of this course is critical for students to improve their programming skills and better prepare them for professional life. In this study, the transcripts of the interviews with teachers were used in the analysis in order to evaluate the outcomes of the "Fundamentals of Programming" course. In this context, the results of human and AI-supported analyses were compared using the constant comparative method, one of the qualitative research methods. The effectiveness of AI-supported analysis in extracting meaningful information from large data sets and recognizing complex patterns is increasing (Giray, 2023).

The significance of the research is that the use of AI-supported tools in qualitative data analysis provides efficiency and accuracy in educational research. Also, the hybrid prompt design developed in this study is designed to improve the accuracy and validity of AI-supported analysis. In this context, the study provides important insight on how AI-supported qualitative data analysis can be used more effectively in educational research.

## **Literature review**

### ***Artificial intelligence***

Artificial intelligence (AI) is one of today's most significant technological advancements, constantly evolving with ongoing efforts to model human intelligence (Göçmez, 2023). A pivotal moment in AI history was the Turing Test, developed in the 1950s to evaluate machine intelligence. Another milestone was the Dartmouth Conference in 1956, where McCarthy introduced the concept of artificial intelligence, laying the foundation for future AI research (McCorduck, 2004). Deep learning, a subset of machine learning, utilizes artificial neural networks to recognize complex patterns in large datasets. It processes information through multiple layers, enabling significant advancements in image recognition, natural language processing, and voice recognition (Inik & Ülker, 2017). Machine learning is a method that uses algorithms to learn from data and can make predictions based on new data. Based on this learning, this process enables models trained with sufficient and high-quality data sets to fulfil real-world tasks through sample cases (Serbest, 2024). Advances in subfields such as deep learning, natural language processing, and image recognition have enhanced AI's ability to perform human-like tasks (Goodfellow, Bengio, & Courville, 2016). Notably, the rise of deep learning has revolutionized the way machines interpret visual and verbal information.

ChatGPT is a chatbot built on Large Language Models (LLMs), a language-based artificial intelligence model developed by OpenAI and released on November 30, 2022 (OPENAI, 2022). LLMs are advanced artificial intelligence models capable of producing human-like text, language understanding, and production in language processing tasks using massive data sets. Developing and training LLMs requires significant computational resources and engineering effort. These models use various training and fine-tuning strategies to improve their performance across a wide range of applications. LLMs have significantly impacted the AI community's approach to developing general-purpose AI systems, with impressive capabilities such as ChatGPT (OpenAI, 2023). However, challenges such as understanding how new capabilities emerge with scale, the cost of training, and aligning model behaviours with human values are topics of active research and discussion. Since the mid-2010s, LLMs, which have played a central role in research in natural language processing (NLP), have made significant technological advances thanks to three main developments (OpenAI, 2024):

- Increased computational power,
- Expansion of quality and organized text datasets
- Development of new algorithms

Many of these innovations are based on deep learning techniques that enable automatic learning from data using complex artificial neural networks. In particular, advances such as recurrent neural networks and attention mechanisms have allowed these models to parse and retain more important parts of texts and ignore less important ones (Henrickson & Meroño-Peñuela, 2023).

### ***AI in educational research***

In recent years, the applicability of AI in education has expanded, with multiple studies exploring its potential. Research by Elbanna and Armstrong (2023) suggests that ChatGPT can be effectively integrated into education to automate routine tasks and enhance student learning. Similarly, Smith et al. (2022) and Johnson (2021) emphasize that AI-driven tools can improve classroom engagement and provide adaptive learning experiences. The integration of AI in education includes case studies examining students' knowledge construction processes, as well as teachers' roles in AI-enhanced classrooms. While these studies demonstrate AI's effectiveness in educational research, they also highlight the complexities and challenges involved in its application. Artificial intelligence (AI) is one of today's most significant technological advancements, constantly evolving with ongoing efforts to model human intelligence (Krstić, Aleksić, & Krstić, 2022).

In the literature review, it is seen that AI is used in different functions in educational research. These are:

- Intelligent Tutoring Systems (ITS) and Adaptive Learning Environments: Intelligent Tutoring Systems (ITS) and Adaptive Learning Environments have the capacity to provide instruction tailored to individual student needs. These systems ease the workload of teachers and improve students' learning processes (Luckin et al., 2016; Dağ, 2022). In particular, GeoQA and Talk Moves

applications demonstrate the power of ITS in solving geometry problems and increasing interaction in mathematics lessons (Dağ, 2022; Göçmez, 2023).

- **Student Modeling and Learner Analytics:** Student Modeling and Learner Analytics personalize instructional strategies by analysing students' performance and learning behaviors. These techniques are critical for monitoring students' progress in the learning process and providing them with appropriate support (Dağ, 2022; Göçmez, 2023; Luckin et al., 2016; Ojha, 2023).
- **Automated Scoring and Feedback:** Automatic Scoring and Feedback reduce teachers' assessment burden and provide quick feedback to students (Ifelebuegu, 2023). These systems help students improve their writing skills by assessing written work objectively and consistently (Dağ, 2022).
- **Intelligent Virtual Agents and Conversational Agents:** Intelligent Virtual Agents and Conversational Agents support learning and increase student motivation by interacting with students in natural language. This technology provides students a personalized learning experience (Luckin et al., 2016).
- **AI-powered Educational Tools and Applications:** AI-supported educational tools and applications offer opportunities for personalization at all levels of education. For example, platforms like iTalk2Learn and Duolingo support math and language learning, offering students individualized learning paths (Dağ, 2022; Stone et al., 2016).
- **AI-generated Content and Curriculum Design:** Artificial intelligence plays an important role in content and curriculum design. AI helps teachers and curriculum developers to prepare content that is suitable for students' needs. In particular, AI has a great contribution in extracting text from latex-based documents and producing course materials suitable for student needs (Dağ, 2022; Luckin et al., 2016).
- **AI for Accessibility and Inclusive Education:** AI can potentially increase accessibility and inclusion in education (Bozkurt, et al., 2023, pp. 60–61; Ifelebuegu at all., 2023). Adaptive learning tools, especially for students with disabilities, facilitate all students' access to educational materials and improve their learning experience (Dağ, 2022; Ifelebuegu, 2023; Göçmez, 2023).

The various functions of AI education have the potential to significantly improve learning and teaching processes by enabling us to understand student needs better and provide appropriate solutions to these needs. However, despite these positive impacts of AI in education, researchers such as Feuston and Brubaker (2021) reveal that the integration of AI technologies, and these technologies raises important issues such as ethics, privacy, and accuracy. In this context, while taking full advantage of AI's potential in education, it is necessary to adopt a careful approach to overcome the challenges and minimize this technology's potential risks. The use of AI in research and education raises issues of ethics, privacy, and accuracy (Christou, 2023a). Therefore, it is emphasized that AI should be carefully integrated into educational and research applications and used under human supervision (Christou, 2023b).

### ***AI in qualitative research***

The integration of artificial intelligence (AI) and qualitative research is emerging as an area where researchers are beginning to reshape their approach to qualitative data and analysis processes (Haque et al., 2022). These technologies, especially AI algorithms, offer improved efficiency in tasks that traditionally require human analysts to spend significant time and resources analysing large datasets (Xiao et al., 2023). Researchers have used computers and other technological tools to support qualitative studies (Sellen, 1995; Thunberg & Arnel, 2022; Weitzman & Miles, 1995). From the 1990s onwards, as the internet became more integrated into people's lives, digital methods for researchers began to be evaluated and discussed (Sellen, 1995). During this period, new methodologies such as 'netnography', which combines the internet and ethnography, were introduced, and the use of digital research methods has gradually increased since the 2000s (Kozinets, 2015). With the Covid-19 pandemic, significant changes in 2020 required researchers to adapt to the new situation, which can be considered a digital revolution (Nind et al., 2021). Digital methods have provided an important alternative when researchers cannot conduct face-to-face interviews due to geographical distance, health issues, or social distancing requirements (Thunberg & Arnel, 2022).

AI models can provide an initial analysis of textual data by summarizing content, identifying key themes, sentiments or trends, and even providing insightful recommendations and generating questions that can guide further research (Cui et al., 2023; Khurana et al., 2023; Rietz & Maedche, 2021 Shaik et al., 2022). Researchers can increase the validity and reliability of their studies by comparing AI's findings with human analysis (Gebreegziabher, 2023). AI's capacity to process data quickly allows researchers to conduct real-time analysis during the data collection phase and adjust research methodologies accordingly based on preliminary findings (Panda et al., 2019). The development of automated qualitative analysis techniques has enabled researchers to analyse volumes of data that they may struggle with manually (Welsh et al., 2002), and the rise of large language models (LLMs) may further improve the efficiency of analysis processes (Zhang et al., 2023).

Especially in data analysis, AI offers a wide range of applications in various processes, such as data preprocessing, classification, clustering..., which allows researchers to perform complex analyses and obtain more accurate results (Davenport & Ronanki, 2018). The use of AI tools such as Big Language Models provides researchers with the opportunity to quickly identify patterns and relationships in data sets while at the same time increasing the objectivity of the analysis and improving the reliability of the results (OpenAI, 2024).

Despite the impressive capabilities of AI, ML and LLMs, the complex nature of qualitative analysis presents unique challenges that these technologies are still trying to overcome (Hämäläinen, Tavas & Kunnari, 2023)

However, it should be noted that AI tools should be used with caution and evaluated by human observers due to potential problems such as "hallucinations" (Athaluri et al., 2023).

### **Prompt engineering**

Prompt engineering is an emerging discipline in artificial intelligence (AI) and natural language processing (NLP). It deals with designing input commands (prompts) to obtain desired outputs from AI, such as language models. The addition of the term "engineering" to this field emphasizes that this process requires a systematic and analytical approach, making it a combination of both creative and technical skills (Giray, 2023). The problem addressed in Prompt engineering is how to communicate effectively with AI to achieve specific results, and this requires a range of skills such as creativity, technical understanding, and iterative improvement (Spasić & Janković, 2023).

Prompt engineering combines elements of both art and science, as a process that aims to achieve desired outputs when interacting with language models (Shao et al., 2023). The heuristic approaches used in this process provide rules based on experience and intuition, allowing for the rapid identification and implementation of effective prompt design strategies (Sivarajkumar et al., 2023). This represents a synergy between understanding how language models work (science) and designing creative and effective prompts (art), making prompt engineering a rich field of research and practice (Spasić & Janković, 2023).

Building on the concept of prompt engineering, previously defined as both an art and a science, hybrid prompt engineering takes this further by integrating the creative, heuristic aspects with scientifically validated principles. While crafting prompts involves intuition, experience, and creativity to coax the best possible responses from AI models, the science behind it relies on empirical research and established guidelines to ensure consistency and reliability. Therefore, Hybrid prompt engineering merges these two dimensions—using heuristic strategies to quickly generate effective prompts while grounding them in proven, research-based techniques. This dual approach aims to leverage the strengths of both worlds, optimizing model performance and ensuring robustness in outcomes. Studies such as those by Perez et al. (2021) highlight how carefully structured prompts significantly enhance AI performance, illustrating the value of combining intuitive methods with scientifically backed practices.

PE is critical to enable AI to produce more accurate and context-appropriate responses, especially in the fields of natural language processing and text analysis (Radford et al., 2019). PE is the practice of creating tailor-made input commands that aim to elicit clear and useful responses from AI systems such as LLM (Petroni et al., 2020). With the rapid advances in AI technology, the ability to have an effective dialog with these systems is becoming increasingly important. This ability, known as PE, is defined as a skill set that enables the rapid achievement of desired results, and expertise in this area is seen as a great asset in technology and engineering.

### **Prompt engineering in qualitative data analysis**

In fields such as qualitative analysis, it is important to use an iterative methodology that continuously adapts and evaluates various prompt engineering strategies to utilize the full potential of ChatGPT (Zhang et al., 2023). A practical, prompt design is decisive in the user's interaction with LLMs (OpenAI, 2023). Ideally, the design of a prompt first requires the user to clearly articulate their goal or problem, increasing the accuracy and relevance of the output the user expects from the model (Wang et al., 2023). The prompt patterns proposed by White et al. (2023) offer valuable insights into how this process can be optimized. These include:

- With the persona pattern allows the user to direct the model to play a specific role, Context Manager empowers the user to manage the contexts that the model should consider or ignore.
- The Flipped Interaction pattern allows the model to become a more active participant when the model needs to ask questions to help the user achieve their goals.
- The Cognitive Verifier pattern encourages more accurate solutions to a problem by breaking it down into sub-questions.
- Alternative Approaches encourage the user to think about and evaluate different solutions.

- Question Refinement helps to obtain more accurate information by making the user's question clearer and more focused.
- Template pattern allows the user to receive information in a specific format or template. The implementation of these patterns constitutes the basic steps of prompt design and allows users to interact more effectively with LLMs (White et al., 2023). Therefore, for an effective prompt design, these patterns should be utilized and customized according to users' goals (Wang et al., 2023).

Zhang et al. (2023) state that prompt design is an important step when interacting with LLMs and that successful management of this process is critical to achieving the desired results and suggest the following steps for a functional prompt:

- For an effective, prompt design, it is first necessary to describe the background and conceptual understanding of the task in detail. This provides the model with context about the task and allows it to develop a more nuanced understanding.
- Focusing on the methodology, the specific task that the model should perform should be clearly stated. Here, it is important to provide guidelines about the analysis process and indicate which methodologies or theories the model should use.
- The format of inputs and outputs should be defined; this gives the model a clear understanding of what data structure it will process and in what format the results should be presented.
- Role-playing allows the model to focus more on completing a specific type of task. At the same time, prioritization aims to focus on the main themes of the analysis and limit the handling of secondary information.
- Transparency and traceability can increase users' confidence in the results produced by the model, while the acceptance of expertise can ensure that the model produces consistent outputs.
- The answer to the question of how a functional prompt for qualitative data analysis should be asked to ChatGPT is as follows:
- Clear and Descriptive: Clearly explain what the data set you want to analyse is about. Provide detailed information about the content and structure of your data set and your expectations from the analysis.
- Specific Questions: Specify which questions you are looking for answers to. This helps to define the focus of the analysis.
- Data Format: Specify the format of your data set (e.g. CSV, Excel, JSON) and the important columns. If possible, explain what the column headings mean.
- Analysis Type: Specify the type of analysis you want to be performed (e.g. trend analysis, correlation research, thematic data analysis).
- Time Period (if applicable): Specify the time period covered by the data. This is especially important for time series analysis.
- Special Requests: Indicate if you have any requests for graphs, tables or any special format. Also, if there are any special criteria or parameters you would like to be taken into account in the analysis, include them (ChatGPT, query dated May.2024).

### Importance of the research

In this study, teachers' opinions were obtained through interviews to evaluate the outcomes of the "Programming Fundamentals" course in vocational high schools. The results were compared with two analyses (human and artificial intelligence) to develop in-depth knowledge and understanding, especially in educational research. This study also uses the constant comparative method, one of the qualitative research methods, to report the processes of analysing the interview texts given to artificial intelligence.

Qualitative research is a widely used method in social sciences and many disciplines. Qualitative data are usually obtained from texts, interviews, observations, and other unstructured sources. Analysing this data can be a time-consuming and exhausting process for researchers. With the development of artificial intelligence technologies, new opportunities in qualitative data analysis have emerged. AI models successfully extract meaningful information from large amounts of data and recognize complex patterns. However, to use these models effectively, it is important to determine appropriate prompts. Due to the nature of qualitative data, developing these prompts can be a challenging process.

This study aim to develop a hybrid (both heuristic and scientific) approach for writing prompts to be used in qualitative data analysis and to compare human researcher and AI-supported results. Heuristic approaches are methods based on heuristic strategies and experience rather than precise rules (Shao, 2023). Such an approach can better address the uncertainties and versatility in the interpretation of qualitative data. The study also incorporates empirically grounded strategies validated through academic research to bridge this with the scientific aspect of prompt engineering. By combining heuristic methods with evidence-based rules, the proposed hybrid approach not only addresses the flexibility required in qualitative data

interpretation but also ensures rigor and consistency in the analysis process. This combination seeks to enhance the robustness and reliability of prompts used in qualitative research, offering a balanced approach that capitalizes on both experience-driven insights and scientifically proven methodologies.

In our research, to fully exploit the potential of AI-supported data analysis in qualitative data analysis, we adopted an iterative methodology involving various prompt engineering strategies, as Zhang et al. (2023) outlined. In this process, we emphasized OpenAI (2023) and Wang et al. (2023), who emphasize that the research question and goals should be clearly stated first in the prompt design. In particular, we created a functional prompt for ChatGPT, inspired by the 'Persona', 'Context Manager', 'Flipped Interaction', 'Cognitive Verifier', 'Alternative Approaches', 'Question Refinement' and 'Template' patterns proposed by White et al (2023).

While preparing this prompt, we positioned ChatGPT as a vocational high school computer teacher and an analyst specializing in qualitative data analysis. This is based on the 'Persona' design. Our prompt design, in which the research topic and data collection method are described in detail, is an application of the 'Context Manager' and 'Question Refinement' patterns. We applied the 'Flipped Interaction' design by asking ChatGPT to take an active role in the analysis process and conduct a detailed analysis using a specific qualitative analysis method, such as the constant comparison method.

In our prompt's testing and revision phases, we used the 'Cognitive Verifier' pattern to verify the accuracy of ChatGPT's outputs and the logical steps of the analysis process. We also used the 'Alternative Approaches' pattern to ask ChatGPT to present alternative methods for analysis and evaluate their advantages.

In conclusion, this prompt design and iterative development process provided important insights into how ChatGPT can be effectively used in qualitative data analysis. Insights from sources such as Zhang et al. (2023), OpenAI (2023), and Wang et al. (2023), and the patterns suggested by White et al. (2023), formed the basis of our prompt design and allowed us to maximize the analysis capabilities of ChatGPT. This process provided a framework for future research, showing how ChatGPT and similar LLMs can be used effectively in qualitative research in educational sciences and beyond.

Identifying appropriate prompts to improve the efficiency of qualitative data analysis can significantly reduce the time and effort of researchers by enabling AI models to produce more accurate and meaningful results. Artificial intelligence can only analyse a data set of this size in a controlled manner will only be possible by entering the correct prompts.

AI-supported analysis can enable more in-depth data analysis, thereby improving the quality and validity of qualitative research. In fact, comparing findings from different perspectives (human researcher vs. AI-supported) provides a broader perspective on the topic. It also ensures transparency of the analysis process, remaining objective when making comparisons and placing the findings in an overall context. In this way, the reliability and validity of our academic work can be enhanced.

In conclusion, this study aims to fill an important gap in the field of qualitative research and promises to provide researchers with a more efficient, deeper, and more collaborative analysis process.

## Research questions

- What kind of a prompt design and optimization approach can be developed for AI-supported qualitative data analysis?
- What are the findings of human researchers and AI-supported analysis regarding the evaluation of the "Fundamentals of Programming" course?
- What are the similarities and differences between human research and AI-supported analysis?

## Method

### Research design

This study used a comparative case study design. Comparative case studies aim to reveal similarities and differences and make comparisons by examining more than one situation in depth (Stake, 2006; Yin, 2003). This design is frequently preferred in social sciences as it enables the comparison of different situations or contexts (Baxter & Jack, 2008; Chmiliar, 2010).

In comparative case studies, the situations examined can be interrelated or independent of each other (Gustafsson, 2017). In this study, a comparative case study design was chosen because examining programming instruction with two different analysis methods (human researcher and artificial intelligence

analysis) requires a comparative perspective. Thus, the similarities and differences between the analysis results can be revealed.

A comparative case study examines a phenomenon across different contexts. This study evaluates the rapid data processing capabilities of AI-powered analysis tools compared to the in-depth analysis capabilities of human researchers by using the constant comparison method in data analysis on the same data set. In comparative case studies, such patterns, themes, similarities, and differences between cases are usually explored.

### **Population and sample**

The participants of this study were 10 ICT teachers working in four different Vocational and Technical Anatolian High Schools in a province located in Central Anatolia. The criterion was that the selected teachers should have at least five years of professional experience, which indicates that the participants have the potential to provide in-depth information and collect comprehensive data. The participants' professional experience ranged from five to twenty-three years, and teachers with at least five years of experience teaching the Fundamentals of Programming course were considered a priority. However, this period was reduced to two years due to insufficient participants.

### **Data collection process**

Data collection was done through face-to-face semi-structured interviews with teachers. The interview form, which was revised based on the opinions of two experts. Moreover, to increase the trustworthiness of the interview form, it was tested in a pilot study to check the clarity and accuracy of the questions and their alignment with the research objectives. The interviews were designed to collect in-depth information about the teaching and learning processes of the course and their relationship with professional competencies.

### **Data analysis method**

Data analysis was carried out in three stages:

- **Human Researcher Qualitative Data Analysis:** The data were analysed using qualitative coding techniques within the constant comparison method of grounded theory. Open, axial, and selective coding were applied to systematically examine and categorize the data. In the open coding phase, initial concepts were identified based on observed similarities and differences. Axial coding then structured these concepts into main categories, while selective coding refined them into a unified theoretical framework. This approach facilitated a clearer understanding of category relationships and ensured a structured analysis.
- **AI-supported Qualitative Data Analysis:** The same data set was analysed using the ChatGPT Data Analyst tool. Similarly, ChatGPT applied open, axial, and selective coding stages with a constant comparison method. Artificial intelligence analysed the data using its ability to summarize content and summarize important information and tried to create a theoretical framework.
- **Comparative Analysis:** As last step for the data analysis of this research, the results of both analysis (human researcher vs AI-supported) were brought together to reveal the problems of programming courses. This way, researchers identify themes and categories and compare the similarities/differences between human researchers and AI-supported results. This comparison aims to develop an in-depth understanding of the advantages, limitations, and potential uses of both methods in a qualitative research context.

### **Limitations**

This study mainly faces technological limitations due to the constraints of AI tools in processing complex and unstructured datasets. AI-based analysis methods may struggle with context-dependent interpretations, potentially affecting the accuracy of findings. Additionally, the study relies on Chatsupported.0, a specific AI tool used between February and April 2024, limiting the results to queries from this timeframe.

## **Findings**

### **What kind of a prompt design and optimization approach can be developed for AI-supported qualitative data analysis?**

In this study, we develop a hybrid prompt design and optimization approach for AI-supported qualitative data analysis. This approach is in line with existing prompt design processes in the literature and is supported by heuristic methods. In light of the information obtained from user experiences and the prompt

design processes proposed in the literature, have been improved or modified. The prompt design process consists of two main phases:

**Phase 1: Preliminary preparation and pilot testing**

Developing and refining the prompt through iterative testing and incorporating literature-based strategies.

1. Pilot Test with Single Interview:

Objective: Experiment with the initial prompt design.

Action: Used an unpublished interview to create a preliminary prompt in colloquial language, incorporating research questions. Based on the iterative approach by Zhang et al. (2023) tested clarity in stating research questions and objectives using ChatGPT's data analysis tool.

2. Prompt Refinement Using 'Question Refinement' Pattern

Objective: Ensure ChatGPT accurately analyses the interview.

Action: Updated the prompt according to White et al. (2023) to focus ChatGPT solely on the interview content. This process is designed to ensure that ChatGPT only operates on the interview text and that the interview text is correctly analysed by ChatGPT.

3. Output Format Update with 'Template' Pattern

Objective: Standardize the presentation of analysis results.

Action: Applied the 'Template' pattern from White et al. (2023) to define a specific format for ChatGPT's outputs. This approach enabled ChatGPT to present the analysis results in a specific format.

4. Functional Evaluation and Heuristic Updates

Objective: Enhance prompt effectiveness through continuous improvement.

Action: This process involves continuous improvement of the prompt. The aim is to assess prompt functionality and make necessary updates following a heuristic approach and iterative methodology (Zhang et al., 2023). Also, getting a consultation from ChatGPT for suggestions on prompt design enhancements.

5. Testing Data Format Compatibility and Ensuring Comprehensive Analysis

Objective: Determine ChatGPT's ability to process and thoroughly analyse entire interview transcripts. This is a step pointed out by Wang et al. (2023), who emphasize the importance of thoroughly analysing the entire data set.

Action: Tested various data formats (Word, PDF, CSV, TXT) for compatibility.

Identified and addressed limitations in data processing.

Redesigned prompts for effective handling of the entire dataset.

Verified thoroughness by asking detailed questions, ensuring no data was overlooked.

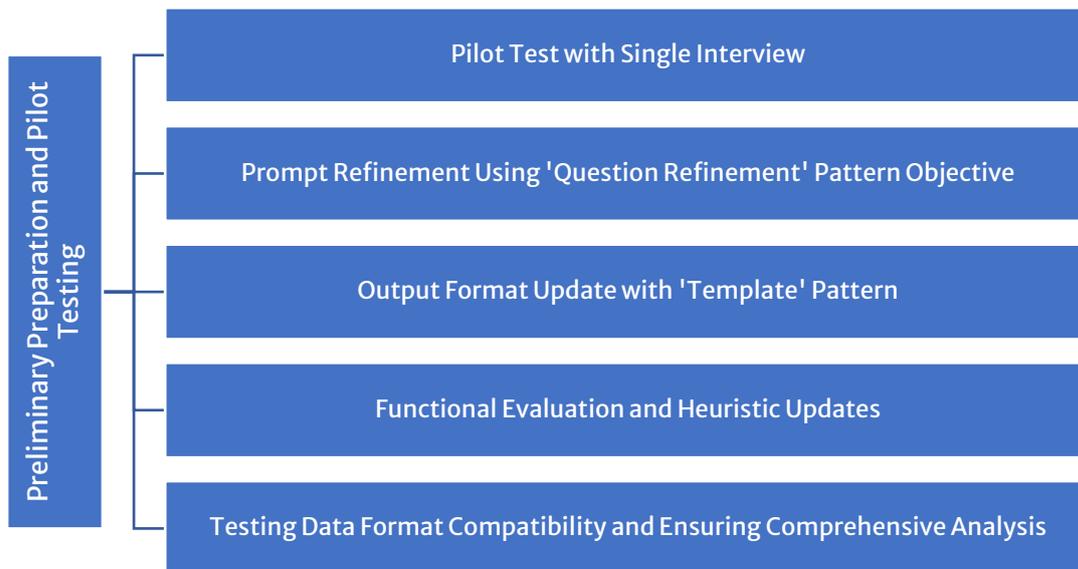


Figure 1. Preliminary preparation and pilot testing scheme

### **Phase 2: Final hybrid prompt design**

This phase involves a three-stage process to enable ChatGPT to perform the constant comparison analysis method effectively like a human researcher.

#### 1. Task Definition and Role Assignment

Objective: Establish context and assign a specific analytical role to ChatGPT.

Prompt Components:

- **Persona:** Instructed ChatGPT to adopt the role of an analyst specialized in qualitative data analysis and instructional technologies, with 12 years of experience as a computer teacher in vocational high schools.
- **Data Type:** Provided details about the semi-structured interviews conducted with 10 teachers, specifying file formats and participant notation for accurate reference.
- **Study Topic:** The research focuses on the relationship between learning-teaching processes and assessment methods in programming fundamentals courses from teachers' perspectives.
- **Research Question:** Clearly state the specific research question to guide the analysis.

Crafted a comprehensive prompt incorporating the above components to set the stage for the analysis. ChatGPT was informed about the task with an explanatory prompt, which can be seen below. Each code section was coloured with a different colour.

Hello, you are an analyst specialized in qualitative data analysis and instructional technologies who spent 12 years as a computer teacher in vocational high schools. You have worked in a vocational high school in Niğde for many years and the experience you have gained during this time may be important for your analysis. (Persona)

The research examines the factors affecting the teaching and learning activities in the Fundamentals of Programming course in Vocational High Schools from the perspective of teachers. The research is planned as a case study aiming to improve the teaching and learning processes of this course. The results of the analyses you will make in this study are planned to be published as an article in an SSCI journal. (Study Topic)

The research question of this study is "What are teachers' views on the factors affecting the Fundamentals of Programming course in Vocational High Schools?". (Research Question)

In the data collection phase, interviews were conducted with information technology teachers teaching the programming basics course using a semi-structured interview form. Interview records made with 10 different teachers will be given to you as 11 separate files. The participant of the interview is indicated with "(( ))". For example ((Participant 2 Interview)). Because when I ask you for quotations about the findings of this study, you need to indicate which participant the quoted part belongs to. The text after "Pınar Arslan Interviewer:" is the interview question; the text after "PARTICIPANT 1:", "PARTICIPANT 2:" etc. is the answer to the question. Teacher 1, the first part of the interview file is attached here. Now I want you to analyse this interview text sentence by sentence and wait for my command to analyse it. (Data Type)

## 2. Data Submission and Examination

Objective: Prepare data for analysis and instruct ChatGPT on the analytical process.

Prompt Components:

- Data Submission:

Divided lengthy interview transcripts into manageable segments.

Submitted each part with clear instructions.

Indicated which part of the teacher's interview was attached.

- Data Examination

Reiterated the request for sentence-by-sentence analysis.

Asked ChatGPT to wait for further commands before analysing.

A separate Word file was created for each of the interview texts. A very long interview text was divided into two and transferred to the data analysis tool of the ChatGPT with separate prompts that can be seen below. It was instructed on how to analyse the data.

The .... part of the teacher interview file is attached to this attachment. (Data submission)

I want you to analyse the text of the interview with 10 different teachers, which I have just given you in 11 parts (Data submission)

Now I want you to analyse this interview text sentence by sentence and wait for my command to analyse it (Data examination)

## 3. Execution of Constant Comparison Analysis

Objective: Guide ChatGPT through the qualitative data analysis process using constant comparison method, structured under Coding and Formatting.

Prompt Components:

I. Coding: The aim is not to use external information, context, or prior knowledge. Analysis should be solely on the participants' responses and findings should be align with the research questions.

- Open Coding: Break down data into categories and subcategories related to the research question.

Instructions:

Create categories with minimal overlap.

Identify key themes emerging from the data.

- Axial Coding: Examine relationships and connections between the categories from open coding.

Instructions:

Explore how categories relate to each other and the research question.

Provide a deeper understanding of the relationships

- Selective Coding: Identify the core theme and integrate all categories around this central concept.

Instructions:

Determine the central theme that addresses the research question.

Organize other categories around this core theme.

II. Formatting: Visual representation of relationships between core theme, categories, subcategories, explanations, quotations.

- Detailed explanation of the core theme, categories, subcategories and their connection to other categories.
- Support core themes, categories, and subcategories with significant quotations, indicating the participant interviews as a source.
- Develop a model, diagram or table to illustrate connections.

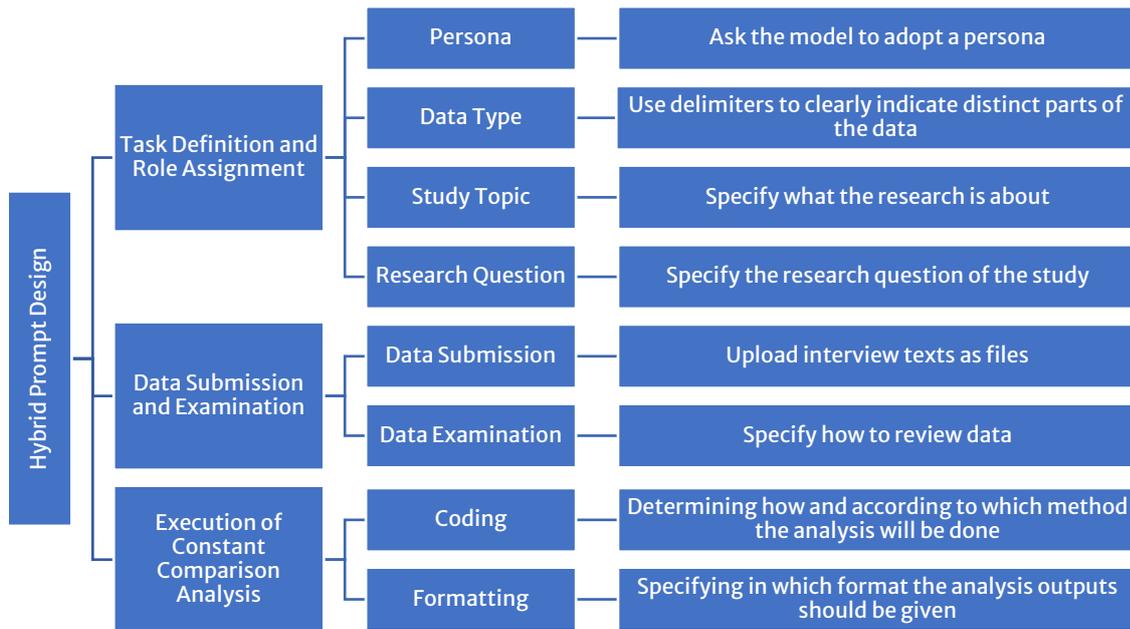


Figure 2. Final hybrid prompt design scheme

In Phase 3, in order for the ChatGPT to perform data analysis, the following prompt design was applied. The aim here is to perform the open axial and selective coding operations, which are the steps of constant comparison analysis, respectively.

Because in this analysis you will use the constant comparison method, which is one of the qualitative data analysis methods. You will not use any additional information, context or your own prior knowledge. When analysing, you only need to act according to the participant responses. The first stage of your analysis is open coding. Open Coding: You will analyse the data obtained from the interview texts by dividing them into categories and subcategories related to the research question. (Coding)

In this process, you will create categories that overlap as little as possible based on the interview and support these categories by quoting important parts. You will present these findings in a table. In the table; Categories should be in the first column, sub-categories in the second column, explanations in the third column and at least 2 quotations in the fourth column. The number of categories should not exceed 5. (Formatting)

- Can you do the open coding analysis for me in the first place? (Coding)

- Thank you, you did open coding successfully. Now it is time for axial coding. At this stage, you will use the categories you obtained in open coding and examine the relationships and connections between these categories. This process will provide a deeper understanding of the relationship between the assessment and evaluation methods of the course and the learning-teaching processes. (Coding)

Could you do the axial coding for me, and I would be grateful if you could develop a model or a diagram that shows the relationship more clearly. (Formatting)

- Now, can you do selective coding using the information you have obtained from the previous stages? (Coding)

By determining the core theme of the research, you should explain how you can organize and relate all other categories around this core theme by supporting it with quotes from the participants (indicating which participant it belongs to). (Formatting)

**What are the findings of human researcher and artificial intelligence analysis regarding the evaluation of "Fundamentals of Programming" course?**

**Comparison of main categories:**

Table 1. AI supported data analysis output

Main Category	Category
Teacher Competencies	Professional Development
	Education Methods

Student Readiness	Interest and Motivation Prejudice and Fears
Course Content and Programming Languages	Curriculum Content Applicability
Measurement and Evaluation Methods	Evaluation Process Measurement of Outcomes
Market Needs and Course Applicability	Market Compliance Professional Applicability

According to AI supported data analysis, the findings related to the "Fundamentals of Programming" course were categorized as follows: The category of teacher competencies includes professional development and training methods. The category of student readiness includes the subcategories of interest and motivation and prejudices and fears. Under the category of course content and programming languages, there are the categories of curriculum content and applicability. Under the category of assessment and evaluation methods, the sub-categories of assessment process and measurement of outcomes are included, while under the heading of market needs and applicability of the course, the categories of market compatibility and vocational applicability are identified.

Table 2. Human analyst output

Main Category	Category
Teacher Competencies	Field Knowledge–Pedagogical competencies (education methods)
	Professional Field Knowledge (Professional Development)
Student Readiness	Student Profile (Prejudices and fears)
	Computer Thinking Skills
	Motivation (Interest and Motivation)
Vocational and Technical Education System and Courses	Vocational High School Entrance System
	Market Needs (Market compatibility)
	Curriculum
	Skills Training (vocational applicability)
Learning and Teaching Process	Measurement and Evaluation

According to the human analyst's analysis, the findings related to the "Fundamentals of Programming" course were categorized as follows: The category of teacher competencies includes content knowledge–pedagogical competencies (teaching methods) and professional content knowledge (professional development). The category of student readiness includes the subcategories of student profile (prejudices and fears), computer thinking skills, and motivation (interest and motivation). Vocational and technical education systems and courses includes the categories of vocational high school entrance system, market needs (market relevance), curriculum, and skills training (vocational applicability). Under the heading of learning and teaching process, there is the category of measurement and evaluation.

**What are the similarities and differences between the findings of human researcher and AI supported data analysis?**

Table 3. Human analyst vs AI Supported main category

Category	Human Analyst	AI Supported
Teacher Qualifications	Pedagogical Competences, Professional Content Knowledge	Professional Development, Education Methods
Student Readiness	Student Profile, Computer Thinking Skills, Motivation	Interest, Motivation, Prejudices And Fears
Measurement And Evaluation	X	Evaluation Process, Measuring Gains
Course Content and Programming Languages	X	Curriculum Content, Applicability
Market Needs and Course Applicability	X	Market Harmonization, Professional Applicability
Vocational And Technical Education System and Courses	Vocational High School Entrance System, Market Needs, Curriculum, Skill Training	X
Learning And Teaching Process	Assessment And Evaluation, Learning Environment	X

According to the findings of human analysts, teacher competencies were identified as pedagogical competencies and professional content knowledge. According to AI-supported data analysis, professional development and training methods are prominent in this category. Human analysts' findings include student profile, computer thinking skills and motivation. According to AI-supported data analysis, student readiness focuses on interest, motivation, prejudices and fears. This category was not mentioned in the human analysts' findings. According to the GPT-4 analysis, the assessment process and measurement of gains are important elements. This category was not mentioned in the findings of human analysts. According to the AI-supported data analysis, the content of the curriculum and applicability are emphasized. This category is not mentioned in the findings of human analysts. According to the AI-supported data analysis, market relevance and vocational applicability are emphasized. Human analysts' findings emphasize the vocational high school entrance system, market needs, curriculum and skills training. According to the GPT-4 analysis, this category was not assessed. Human analysts' findings focus on assessment and evaluation and the learning environment. According to the GPT-4 analysis, this category is not assessed.

**Comparison of subcategories**

Table 4. Human analyst vs AI supported data analysis subcategory

Subcategories	Human Analyst	AI Supported
Pedagogical Qualifications / Education Methods	✓	✓
Professional Field Knowledge/ Professional Development	✓	✓
Curriculum	✓	✓
Interest And Motivation	✓	✓
Market Compliance/Needs	✓	✓
Computer Thinking Skills	✓	X
Student Profile	✓	X
Prejudice And Fears	X	✓
Vocational Applicability/Skills Training	✓	✓
Learning Environment/ Applicability	✓	✓
Measurement Of Outcomes	✓	X
Vocational High School Entrance System	X	✓
Assessment Process / Measurement and Evaluation	✓	✓

In both AI-supported data analysis and human analysts' analysis, there are subcategories of pedagogical competencies/training methods, professional content knowledge/professional development, curriculum, interest and motivation, market relevance/needs, professional applicability/skills training, learning environment/applicability and assessment process/assessment. The subcategories of computer thinking skills and student profile are only present in the human analysts' analysis, but not in the AI-supported data analysis. The subcategories of prejudices and fears and vocational high school entrance system are included only in the AI supported data analysis, but not in the analysis of human analysts. The sub-category of measurement of achievements is included only in the analysis of human analysts, but not in the AI-supported data analysis.

**Selective coding**

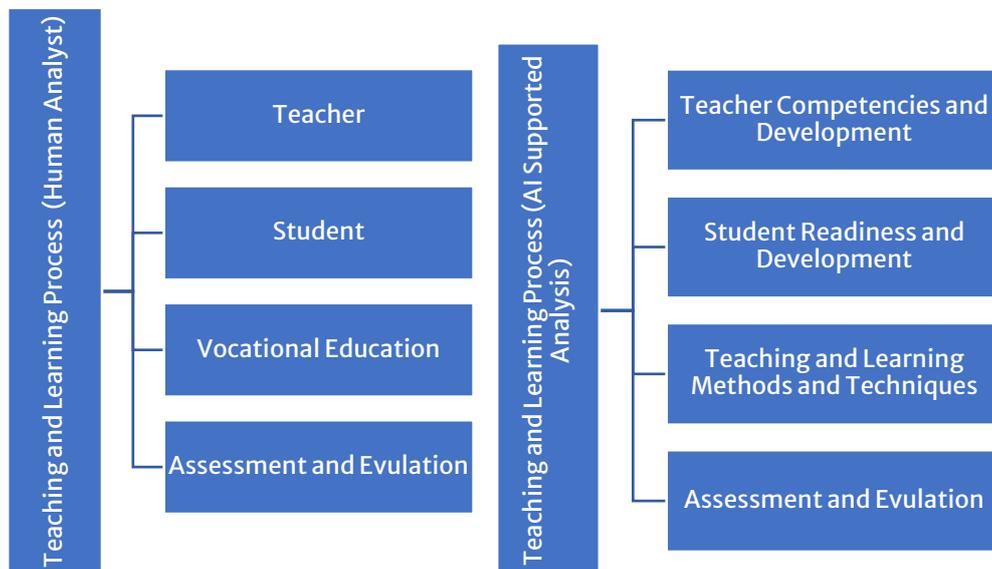


Figure 3. Human analyst vs AI supported data analysis selective coding categories

Figure 3 compares the outcomes of AI-assisted data analysis with those of human analysts for specific coding categories. The "Teaching and Learning Process" was broken down into four primary areas by the human analyst: "Teacher," "Student," "Professional Training," and "Assessment and Measurement." From a wider angle, these categories discuss the connections between the various players and procedures that make up the teaching process. In AI-supported analysis, the same process is examined under the following more focused subcategories: "Assessment and Measurement," "Teacher Competencies and Development," "Student Preparation and Development," and "Teaching and Learning Methods and Techniques." This arrangement implies that AI concentrates on more complex and organized categories.

## Discussion and conclusion

### What kind of a prompt design and optimization approach can be developed for AI-supported qualitative data analysis?

In this study, we develop a hybrid prompt design and optimization approach for AI-supported qualitative data analysis. This approach is in line with prompt design processes in the existing literature and is supported by heuristic methods. Based on the information obtained from user experiences, the prompt design processes proposed in the literature have been improved or modified.

First, inspired by the iterative methodology approach proposed by Zhang et al. (2023), we experimented with prompt design on a single interview as a pilot study. Using ChatGPT's data analysis tool, the research questions and objectives were clearly stated, and the feasibility of the analysis was first tested.

In the second stage, the 'Question Refinement' design proposed by White et al. (2023) was applied to update the prompt and ensure that ChatGPT analysed the interview transcript correctly. This process enabled ChatGPT to operate only on the interview text and make a more accurate analysis.

In the third step, the output format was updated using the 'Template' pattern proposed by White et al. (2023), and ChatGPT was enabled to present the analysis results in a specific format. In the fourth step, the functionality of the developed prompt was evaluated, and updates were made when necessary, in accordance with the heuristic approach. By the iterative methodology proposed by Zhang et al. (2023), continuous improvement of the prompt was ensured. Furthermore, in the design of the prompt, ChatGPT itself was asked how the prompt should be designed.

In the fifth step, the dataset was presented in different formats (Word, PDF, CSV, TXT) so that the full interview transcripts could be analysed by ChatGPT. This was done in order to evaluate the analysis capabilities of ChatGPT by allowing the processing of different data formats as recommended by OpenAI (2023). The sixth step checked whether the full interview transcripts had been analysed and focused on how ChatGPT should process the entire dataset. In this process, the changes on the Prompt were realized with the support of ChatGPT.

In the seventh step, the prompt design for constant comparison analysis was divided into three stages. First, ChatGPT was informed about the task with an explanatory prompt, then a separate Word file was created for each of the interview texts and how to analyse the data was specified with a command. In the last stage, ChatGPT was directed to perform open, axial and selective coding operations, which are the steps of constant comparison analysis, in order to perform data analysis.

Kirsten et al. (2024) examined the performance of large language models in qualitative data analysis processes and the comparison of these processes with human analysts. This study shows that big language models have some limitations compared to human analysts but can be quite effective in certain tasks. Similarly, tools such as Prophet (Yu et al., 2024) suggest that large language models can achieve more accurate results by using response heuristics to improve response accuracy.

As a result, the prompt design and optimization process developed in this study improves the existing approaches in the literature. It provides a more effective and efficient method for AI-supported qualitative data analysis. The methodologies of Zhang et al. (2023) and White et al. (2023) formed the basis of this process and were modified with improvements based on user experiences. Considering the findings of Kirsten et al. (2024) and Yu et al. (2024), this approach is considered as an important step to maximize the potential of AI in qualitative data analysis.

– What are the findings of human researchers and AI-supported analysis regarding the evaluation of the "Fundamentals of Programming" course?

The findings regarding the evaluation of the "Fundamentals of Programming" course in vocational high schools showed various similarities and differences between human analysis and AI-supported analysis. While the findings of human analysts focused on issues such as pedagogical competencies, professional field knowledge, and student profile, AI-supported analysis focused more on professional development, training methods, interest, and motivation. Both analytical methods reached similar conclusions on teaching processes, market needs and course applicability. However, it was observed that human analysts were able to make more detailed and contextual comments, while artificial intelligence stood out with its ability to analyse large data sets faster. In this process, the steps of carefully examining the data, identifying meaningful codes and organizing these codes under categories were followed. This method enabled the systematic categorization of the problems faced by the teachers and the solutions they suggested and contributed to achieving the purpose of the study.

In this qualitative data analysis process, human analysts and AI-supported were observed to follow similar methodological steps. As human analysts, in-depth information was gathered through semi-structured interviews, and the data was read carefully to identify recurring themes and patterns. Using constant comparison, relationships were established between the data and categories were created. The context and nuances of the data were understood, enabling a deeper interpretation of themes and categories. During the analysis process, care was taken to minimize personal bias and preserve the natural meaning of the data. The persona content in the prompt used in this study was added to enable ChatGPT to behave like a human analyst. Specifically, ChatGPT was asked to act as "an analyst who has worked as a computer teacher in vocational high schools for 12 years and is an expert in qualitative data analysis and instructional technologies". This persona was designed to enable ChatGPT to use his/her professional experience and contextual knowledge to conduct more in-depth and meaningful analysis like a human analyst.

ChatGPT made sense of the texts by analysing a large dataset and using natural language processing techniques. The data was carefully analysed to identify recurring themes and patterns and coded. It was observed that ChatGPT has the advantage of being able to analyse large amounts of data in a short period of time thanks to its fast access to a large data pool and powerful language processing capabilities. However, it was found that it sometimes had difficulties in understanding the contextual and cultural nuances in depth, and that it could miss some important details in the data analysis process.

Both methods follow a systematic approach, and as human analysts, we were found to be more advantageous in understanding the contextual and cultural meanings of the data in depth. The fact that human analysts focus on subcategories such as student profiles and computer thinking skills shows that they are more advantageous than ChatGPT in understanding the contextual and cultural meanings of the data in depth. Especially in subcategories such as measurement of learning outcomes and teacher competencies, human analysts can conduct more in-depth analysis of the effectiveness of teaching processes and evaluation of student achievement. Human analysts' experience and intuition better capture contextual and cultural nuances that were not captured in ChatGPT's analysis. Moreover, human analysts excel in assessing the complex relationships and cultural fit between education and cultural fit between education and the labour market. These findings support that human analysts have an advantage over ChatGPT in better understanding and evaluating the contextual and cultural meanings of data. ChatGPT, on the other hand, stands out with its ability to perform fast and comprehensive analysis. In an ideal situation, a combination of both methods could increase the accuracy and scope of qualitative data analysis.

- What are the similarities and differences between human researcher and AI-supported analysis?

Studies in the literature compare AI performance in qualitative analysis processes with human analysts. Bahrini et al. (2023) reported that ChatGPT provides speed and efficiency in thematic analysis but needs customized instructions. Teubner et al. (2023) found that AI can be faster and more efficient than human analysts in some stages of thematic analysis, but human analysis is more reliable in terms of accuracy and depth of results. Tai et al. (2024) found that large language models (LLM) can support traditional qualitative coding and provide a systematic and reliable platform for code detection.

By comparing the analysis results of the human analyst and ChatGPT-4, similarities and differences were identified. This comparison covered the findings of both analyses in main and sub-categories. It was examined how the findings of the human analysts, such as pedagogical competencies, professional domain knowledge, student profile and computer thinking skills, overlapped with the categories in the ChatGPT-4 analysis, such as interest, motivation, prejudices and fears, market compatibility and professional applicability.

Both human analysts and AI have strengths and weaknesses in AI-supported qualitative data analysis. The similarities between the findings of human research and AI analysis are that both methods emphasize the importance of teaching processes and market needs. Both analyses found similar results on training methods and teacher competencies. However, the AI analysis saved time by processing larger datasets

quickly and demonstrated the ability to identify complex patterns that may have previously been overlooked. This suggests that AI-powered analytics can enhance efficiency by reducing the workload of human analysts, particularly for large-scale datasets.

However, differences arise when human analysts focus more on specific topics such as computer thinking skills and student profiles. AI analysis focused on more general themes such as prejudices and fears, while human analysts focused on more detailed and specific issues. It was also found that some sub-categories were missing in AI analysis, but in general, the ability to analyse broadly and quickly complemented human analysts' ability to provide in-depth analysis and contextual information.

The AI-supported analysis appears superior to human analysis in quickly and effectively analysing large-scale text-based datasets. For example, Yardley et al. (2022) showed that machine-assisted topic analysis (MATA) took only 40 hours to identify similar main themes, compared to 147.5 hours for human analysts. However, human analysts were found to have an advantage in terms of a deeper understanding of contextual and cultural meanings. This can be explained by the fact that human analysts can better capture the nuances of the data through their experience and intuition, especially in subcategories such as student profiles and computer thinking skills. Moreover, in areas such as the measurement of outcomes and teacher competencies, human analysts' ability to assess complex and multidimensional relationships in educational processes provides an advantage over AI's inability to capture such contextual nuances fully.

However, there is also evidence that AI struggles to capture some complex and contextual nuances. A study by Thomas and Phillips (2020) found that AI-based analysis tools do not fully reflect the contextual knowledge and experience of human analysts, and in some cases can provide misleading results. Especially in qualitative analysis that requires in-depth and fine-grained details, there may be cases where AI is insufficient. This highlights the limitations of AI in understanding context and analysing complex human interactions.

In addition, research by Smith et al. (2021) shows that machine learning algorithms are successful in qualitative data analysis in terms of consistency and accuracy but may miss certain contextual factors. AI may struggle to capture fine details based on the experience and intuition of human analysts. This suggests that AI-based analytics may not always provide the deep contextual understanding and detail that human analysts provide.

In conclusion, AI-supported qualitative data analysis methods can be valuable for rapid and large-scale data analysis in qualitative research. However, they cannot wholly replace human analysts' analysis based on experience and contextual knowledge. The combination of AI and human analysts can improve the efficiency and accuracy of qualitative data analysis. This combination can lead to more comprehensive and reliable results in qualitative research because the speed and efficiency advantages of AI can be complemented by the in-depth understanding and contextually informed analysis capabilities of human analysts. By enabling qualitative research to take a broader perspective, AI-supported analysis methods can improve research quality and lay an important foundation for the wider adoption and integration of AI technologies in education.

## **Recommendations**

- AI and Human Analyst Collaboration: There should be more effective collaboration between AI-supported tools and human researchers. Combining the strengths of human experience and contextual understanding with the speed and efficiency of AI technology can lead to more comprehensive and accurate analysis. In future research, the hybrid use of human researchers and AI-supported tools can enable both in-depth analysis and faster data processing, enhancing the overall quality of qualitative data analysis. This approach allows researchers to leverage the advantages of AI while maintaining the nuanced insight that only human analysts can provide.

- More training and professional development opportunities for teachers and researchers on the use of AI and data analysis. This will enable educators to use AI technologies more effectively and encourage more widespread use of these technologies in classroom practice and research projects.

- Research and Development: Continuous research and development of AI applications in the field of educational technology should be encouraged. By developing new AI algorithms and tools, more value can be added to educational processes and existing applications can be further improved. This can generate innovative solutions in education, enriching teaching and learning experiences.

- Prompt Design and Optimization: In order to use AI tools efficiently, the focus should be on prompt design and optimization processes. These processes should be continuously updated in line with the research questions and objectives. In particular, designing the prompts used in the context of education and training

in accordance with the needs of educators and students will ensure maximum utilization of the opportunities offered by artificial intelligence.

These recommendations provide strategies for a more effective and efficient use of AI technologies in education and aim to maximize the potential of these technologies.

## Declarations

### Ethics statements

Ethical approval was not required for this study as it involved an examination of artificial intelligence tools rather than human subjects.

### Informed consent

This study used an existing dataset for analysis, and therefore no informed consent was required from participants. The dataset used did not contain any identifying personal information, such as images or other sensitive data.

### Availability of data and materials

The data that support the findings of this study are available upon request from the corresponding author. Due to privacy concerns, the data are not publicly available but may be shared with researchers who meet the criteria for access to confidential data.

### Competing interests

The authors declare that they have no financial or non-financial competing interests.

### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Authors' contributions

Nurşin Akman and Pınar Arslan; conceived the study, performed the data collection, and conducted the human analysis. Erman Uzun contributed to the design of the hybrid prompt, AI-assisted analysis, and comparative evaluation. Both authors contributed to writing and editing the manuscript.

## References

- Baxter, P., & Jack, S. (2008). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*, 13 (4), 544–559.
- Bozkurt, A., Karadeniz, A., Baneres, D., Guerrero-Roldán, A.E. & Rodríguez, M.E. (2021). Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century. *Sustainability*, 13 (800), 2–16. <https://doi.org/10.3390/su13020800>
- Chmiliar, L. (2010). Multiple-case designs. In *Encyclopedia of Case Study Research* (pp. 582–584).
- Christou, P. A. (2023). A critical perspective over whether and how to acknowledge the use of artificial intelligence (AI) in qualitative studies. *The Qualitative Report*, 28 (7), 1981–1991. <https://doi.org/10.46743/2160-3715/2023.6407>
- Christou, P. A. (2023a). The use of artificial intelligence (AI) in qualitative research for theory development. <https://doi.org/10.46743/2160-3715/2023.6536>
- Christou, P. A. (2023b). How to use artificial intelligence (AI) as a resource, methodological and analysis tool in qualitative research?. *The Qualitative Report*, 28 (7), 1968–1980.
- Coffey, A., Beverley, H., & Paul, A. (1996). Qualitative data analysis: Technologies and representations. *Sociological Research Online*, 1 (1), 80–91. <https://doi.org/10.5153/sro>.
- Cui, J., Wang, Z., Ho, S.-B., & Cambria, E. (2023). Survey on sentiment analysis: Evolution of research methods and topics. *Artificial Intelligence Review*, 56, 8469–8510. <https://doi.org/10.1007/s10462-022-10386-z>
- Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, 96 (1), 108–116. <https://doi.org/10.1109/ICEST58410.2023.10187269>
- Feuston, J. L., & Brubaker, J. R. (2021). Putting tools in their place: The role of time and perspective in human-AI collaboration for qualitative analysis. *Proceedings of the ACM on Human-Computer Interaction*, 5 (CSCW2), 1–25. <https://doi.org/10.1145/3479856>

- Gebreegiabher, S. A., Zhang, Z., Tang, X., Meng, Y., Glassman, E. L., & Li, T. J. (2023). PaTAT: Human-AI collaborative qualitative coding with explainable interactive rule synthesis. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 362, 19 pages. <https://doi.org/10.1145/3544548.3581352>
- Giray, L. (2023). Prompt engineering with ChatGPT: A guide for academic writers. *Annals of Biomedical Engineering*, 51(12), 2629–2633. <https://doi.org/10.1007/s10439-023-03272-4>
- Glaser, B.G., & Strauss, A.L. (1967). *The discovery of grounded theory*. Aldine.
- Göçmez, L., & Okur, M. R. (2023). Artificial intelligence applications in open and distance education: A systematic review of the articles (2007–2021). *Asian Journal of Distance Education*, 18 (1). <https://www.asianjde.com/ojs/index.php/AsianJDE/article/view/665>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Gustafsson, J. (2017). Single case studies vs. multiple case studies: A comparative study.
- Hämäläinen, P., Tavast, M., & Kunnari, A. (2023). Evaluating large language models in generating synthetic HCI research data: A case study. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) (CHI '23). Association for Computing Machinery, New York, NY, USA, Article 433, 19 pages. <https://doi.org/10.1145/3544548.3580688>
- Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). "I think this is the most disruptive technology": Exploring sentiments of ChatGPT early adopters using Twitter data. *arXiv preprint arXiv:2212.05856*. <https://arxiv.org/abs/2212.05856>
- Henrickson, L., & Meroño-Peñuela, A. (2023). Prompting meaning: A hermeneutic approach to optimising prompt engineering with ChatGPT. *AI & SOCIETY*, 1–16. <https://doi.org/10.1007/s00146-023-01752-8>
- Hwang, G.-J. (2019). Definition, framework and research issues of smart learning environments—a context-aware ubiquitous learning perspective. *Smart Learning Environments*, 6 (1), 1–14. <https://doi.org/10.1186/s40561-014-0004-5>
- Ifelebuegu, A. (2023). Rethinking online assessment strategies: Authenticity versus AI chatbot intervention. *Journal of Applied Learning and Teaching*, 6 (2). <https://doi.org/10.37074/jalt.2023.6.2.2>
- Ifelebuegu, A. O., Kulume, P., & Cherukut, P. (2023). Chatbots and AI in education (AIED) tools: The good, the bad, and the ugly. *Journal of Applied Learning and Teaching*, 6 (2). <https://doi.org/10.37074/jalt.2023.6.2.29>
- İnik, Ö., & Ülker, E. (2017). Deep Learning and Deep Learning Models Used in Image Analysis. *Gaziosmanpasa Journal of Scientific Research*, 6 (3), 85–104. [https://dergipark.org.tr/tr/pub/gbad/issue/31228/330663#article\\_cite](https://dergipark.org.tr/tr/pub/gbad/issue/31228/330663#article_cite)
- Joshi, S., Rambola, R. K., & Churi, P. (2021). Evaluating artificial intelligence in education for next generation. In *Journal of Physics: Conference Series*, 1714 (1), p. 012039. IOP Publishing. <https://doi.org/10.1088/1742-6596/1714/1/012039>
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82 (3), 3713–3744. <https://doi.org/10.1007/s11042-022-13428-4>
- Krstić, L., Aleksić, V., & Krstić, M. (2022). Artificial Intelligence in Education: A Review. *Technics and Informatics in Education – TIE 2022*, 223–228. <https://doi.org/10.46793/TIE22.223K>
- Lee, A. V. Y. (2020). Artificial Intelligence in Education (AIED). *Conference: 28th International Conference on Computers in Education*.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education, London. <https://discovery.ucl.ac.uk/id/eprint/1475756/>
- McCorduck, P., & Cfe, C. (2004). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. AK Peters/CRC Press. <https://doi.org/10.1201/9780429258985>
- Mesec, B. (2023). The language model of artificial intelligence ChatGPT: A tool of qualitative analysis of texts. *Authorea Preprints*.
- Naveed, H., Khan, A. U., Qiu, S., Saqib, M., Anwar, S., Usman, M., ... & Mian, A. (2023). A comprehensive overview of large language models. *arXiv preprint arXiv:2307.06435*. <https://ar5iv.labs.arxiv.org/html/2307.06435>
- Ojha, S., et al. (2023). From robots to books: An introduction to smart applications of AI in education (AIED). *arXiv preprint arXiv:2301.10026*. <https://arxiv.org/pdf/2301.10026.pdf>

- Panda, G., Upadhyay, A. K., & Khandelwal, K. (2019). Artificial intelligence: A strategic disruption in public relations. *Journal of Creative Communications*, 14 (3), 196–213. <https://doi.org/10.1177/0973258619866585>
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.
- Petroni, F., Lewis, P., Piktus, A., Rocktäschel, T., Wu, Y., Miller, A. H., & Riedel, S. (2020). How context affects language models' factual predictions. *arXiv preprint arXiv:2005.04611*. <https://arxiv.org/abs/2005.04611>
- Radford, A., et al. (2019). Language models are unsupervised multitask learners. *OpenAI Blog*, 1\*(8). <https://insightcivic.s3.us-east-1.amazonaws.com/language-models.pdf>
- Rietz, T., & Maedche, A. (2021). Cody: An AI-based system to semi-automate coding for qualitative research. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 394, 14 pages. <https://doi.org/10.1145/3411764.3445591>
- Serbest, A. B. (2024). Demand Forecasting And Inventory Management With Machine Learning. *Unpublished master's thesis*. Bursa Uludağ University Graduate School of Natural and Applied Sciences <https://acikerisim.uludag.edu.tr/items/52611184-0b35-4c90-bd19-4dac173ffd80>
- Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P., & Galligan, L. (2022). A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. *IEEE Access*, 10, 56720–56739. <https://doi.org/10.1109/ACCESS.2022.3177752>
- Shao, Z., Yu, Z., Wang, M., & Yu, J. (2023). Prompting large language models with answer heuristics for knowledge-based visual question answering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14974–14983). <https://doi.org/10.1109/CVPR52729.2023.01438>
- Sivarajkumar, S., Kelley, M., Samolyk–Mazzanti, A., Visweswaran, S., & Wang, Y. (2023). An empirical evaluation of prompting strategies for large language models in zero-shot clinical natural language processing. *arXiv preprint arXiv:2309.08008*.
- Spasić, A. J., & Janković, D. S. (2023, June). Using ChatGPT standard prompt engineering techniques in lesson preparation: Role, instructions and seed-word prompts. In *2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)* (pp. 47–50). IEEE. <https://doi.org/10.1109/ICEST58410.2023.10187269>
- Spasić, A. J., & Janković, D. S. (2023, June). Using ChatGPT standard prompt engineering techniques in lesson preparation: Role, instructions and seed-word prompts. In *2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)* (pp. 47–50). IEEE.
- Stake, R.E. (2006). *Multiple case study analysis*. Guilford Press.
- Tai, R. H., Bentley, L. R., Xia, X., Sitt, J. M., Fankhauser, S. C., Chicas–Mosier, A. M., & Monteith, B. G. (2024). An examination of the use of large language models to aid analysis of textual data. *International Journal of Qualitative Methods*, 23, 16094069241231168. <https://doi.org/10.1177/16094069241231168>
- Thunberg, S., & Arnell, L. (2022). Pioneering the use of technologies in qualitative research—A research review of the use of digital interviews. *International Journal of Social Research Methodology*, 25 (6), 757–768. <https://doi.org/10.1080/13645579.2021.1935565>
- Wang, X., Attal, M. I., Rafiq, U., & Hubner–Benz, S. (2022, June). Turning large language models into AI assistants for startups using prompt patterns. In *International Conference on Agile Software Development* (pp. 192–200). Cham: Springer Nature Switzerland. [https://doi.org/10.1007/978-3-031-48550-3\\_19](https://doi.org/10.1007/978-3-031-48550-3_19)
- Weitzman, E., & Miles, M. B. (1995). *Computer programs for qualitative data analysis*. SAGE Publications. <https://books.google.com/books?id=E4Y5DQAAQBAJ>
- Welsh, E., et al. (2002). Dealing with data: Using NVivo in the qualitative data analysis process. In *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 3 (2). <https://doi.org/10.17169/fqs-3.2.865>
- White, J., et al. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. <https://arxiv.org/abs/2302.11382>
- Xiao, Z., Yuan, X., Liao, Q. V., Abdelghani, R., & Oudeyer, P.–Y. (2023). Supporting qualitative analysis with large language models: Combining codebook with GPT–3 for deductive coding. In *Companion Proceedings of the 28th International Conference on Intelligent User Interfaces* (Sydney, NSW, Australia) (IUI '23 Companion). Association for Computing Machinery, New York, NY, USA, 75–78. <https://doi.org/10.1145/3581754.3584136>
- Yin, R. K. (2003). *Case study research: Design and methods*. Sage Publications.

- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021, 1-18. <https://doi.org/10.1155/2021/8812542>
- Zhang, H., Wu, C., Xie, J., Kim, C., & Carroll, J. M. (2023). QualiGPT: GPT as an easy-to-use tool for qualitative coding. *arXiv preprint arXiv:2310.07061*.