A Comparative Phenomenological Analysis of Teachers' Experiences Regarding Bilingual Students' Educational Challenges Using Colaizzi's Method and GPT-4's Chain-of-Thought Reasoning

Azad Allahkarami — Assistant Professor, Department of Educational Sciences, Farhangian University, Tehran, Iran. E-mail: A.allahkarami@cfu.ac.ir

Fariba Dortaj * — *Corresponding Author*, Assistant Professor, Department of Educational Sciences, Payame Noor University, Tehran, Iran. E-mail: F.D@pnu.ac.ir

ORCID

Azad Allahkarami **http://orcid.org/0000-0001-8623-6281**

Fariba Dortaj **https://orcid.org/0009-0005-0431-5370**

Abstract

The present study aimed to investigate the challenges of bilingual students and to compare human and machine analyses of first-grade teachers' experiences in this context. The research employed a qualitative approach using Colaizzi's descriptive phenomenological method. Data were collected through semi-structured interviews with 15 first-grade teachers from bilingual regions in Uramanat, Iran. The data were analyzed through two parallel methods: (1) human analysis based on Colaizzi's seven-step framework, and (2) machine analysis utilizing GPT-4 and the chain-of-thought prompting technique. The human analysis identified four major categories: educational, linguistic, psychological, and social challenges. In contrast, the GPT-4 analysis classified the data into six themes: fundamental linguistic difficulties, negative psychological experiences, cultural alienation, home-school mismatch, educational disengagement, and lack of institutional support. Comparative analysis revealed significant overlaps between the two approaches, though the human analysis, with its deeper sensitivity to cultural context and conceptual integration, provided a more unified interpretation. Conversely, GPT-4 demonstrated high precision in detecting textual patterns and independent

themes, thereby unveiling novel dimensions of the problem. The findings suggest that when machine-generated outputs are reviewed by human analysts and evaluated within a theoretical framework, the combination can serve as a valuable complement; otherwise, there is a risk of superficial interpretations.

Keywords: Mother tongue, Educational challenges, Phenomenological analysis, GPT-4

Cite this Article: Allahkarami, A., & Dortaj, F. (2026). A Comparative Phenomenological Analysis of Teachers' Experiences Regarding Bilingual Students' Educational Challenges Using Colaizzi's Method and GPT-4's Chain-of-Thought Reasoning. *Educational Measurement*, 17(63), 1-35. 10.22054/jem.2025.86707.3631

Extended Abstract

Introduction

In multilingual contexts where formal education is delivered in a language different from students' mother tongues, linguistic gaps become central obstacles to educational equity (UNESCO, 2021). This issue is particularly pronounced in regions like Oramanat in western Iran, where Kurdish-speaking children begin formal schooling in Persian—the national language—without systematic support for bilingual education. Numerous studies highlight the critical role of mother tongue instruction in early conceptual development, emotional security, and academic identity (Cummins, 2000; Baker, 2008; Bialystok, 2007). When instruction is misaligned with students' home language and culture, it can lead to cognitive overload, emotional withdrawal, and long-term disengagement from schooling (Gay & Hanley, 1999).

While traditional qualitative approaches have documented these challenges, recent advancements in artificial intelligence (AI), particularly large language models like GPT-4, present new possibilities for analyzing qualitative data (Bender et al., 2021; Birhane et al., 2022). These models demonstrate abilities in semantic clustering, pattern recognition, and text interpretation, but questions remain about their cultural sensitivity, interpretive depth, and methodological validity.

Research Questions

- What are the primary problems faced by bilingual students from the perspective of their teachers?
- What are the similarities and differences between the analysis of bilingual students' problems using Colaizzi's phenomenological method and the analysis of the same data performed by the GPT-4 language model employing the chain-of-thought strategy?

Literature Review

2.1. Bilingualism and Educational Equity

Mother tongue-based instruction is consistently shown to enhance early literacy, conceptual learning, and emotional stability in bilingual students (Ball, 2010). Cummins' interdependence hypothesis (2000) argues that proficiency in the first language supports second-language acquisition and cognitive development. Vygotsky's sociocultural theory (1978) further suggests that language mediates learning; when students are denied access to their linguistic tools, their ability to process new concepts is impaired.

In the Iranian context, researchers such as Rahimabadi (2004), Kazemi (2007), and Allahkarami (2018) have documented how the lack of institutional support for linguistic minorities leads to dropout, underachievement, and identity fragmentation. The centralized curriculum and Persian-only instruction exacerbate disparities in linguistically diverse regions like Kermanshah.

2.2. AI and Qualitative Data Analysis

Large Language Models (LLMs) such as GPT-4 offer new capabilities in qualitative analysis, including thematic extraction, semantic clustering, and conceptual summarization (Isoaho et al., 2021). Unlike traditional tools like NVivo or Dedoose, which support human-led coding, GPT-4 can autonomously interpret text through prompt engineering (Wei et al., 2022). Yet, concerns remain about model bias, context sensitivity, and theoretical grounding (Weidinger et al., 2021; Chen et al., 2023). Recent comparative studies (e.g., Gamieldien et al., 2023; Siiman et al., 2023) indicate that while GPT-4 performs

well in surface-level pattern detection, it struggles with contextual interpretation unless guided by well-structured prompts.

Methodology

This qualitative study used a dual-method approach. In the first phase, 15 first-grade teachers in the Oramanat region were interviewed using a semi-structured guide. Interviews focused on identifying major challenges bilingual students face in academic, emotional, and social domains. Colaizzi's seven-step method was used to analyze transcribed interviews manually. In the second phase, the same interview texts were uploaded into GPT-4. A thought-chain prompt was crafted to simulate a phenomenological researcher following Colaizzi's stages. GPT-4's outputs were analyzed for thematic coverage, conceptual depth, and semantic overlap with human-derived themes.

Results

Human analysis identified four major themes: educational difficulties, psychological distress, social isolation, and linguistic barriers. Teachers emphasized how language mismatch affects comprehension, emotional safety, peer relationships, and academic participation. GPT-4 analysis generated six themes: basic linguistic barriers, emotional strain from linguistic pressure, cultural alienation, home-school mismatch, academic withdrawal, and lack of institutional support. Two novel categories (withdrawal and institutional neglect) were absent in the human analysis but surfaced in the machine analysis.

Discussion

Both analyses confirmed the salience of language-related obstacles and emotional consequences. GPT-4 demonstrated competence in recognizing recurrent linguistic and psychological patterns but lacked cultural-contextual nuance. For instance, it interpreted code-switching as an error rather than a cultural-linguistic adaptation. While human analysis benefitted from theoretical frameworks (e.g., CALP, sociocultural theory) , GPT-4's analysis remained descriptive. Yet, the AI's identification of systemic neglect suggests its potential for

high-level thematic abstraction. Informal Turing testing with blind expert reviewers showed that in 80% of cases, they could not distinguish between human and AI-generated themes.

Conclusion

The study illustrates both the promise and limitations of AI in qualitative educational research. While GPT-4 can reliably extract and categorize themes, its insights should be interpreted with caution. A hybrid model (where AI supports data management and initial coding, and human researchers provide theoretical interpretation) can enhance both efficiency and interpretive depth in analyzing complex socio-educational phenomena.

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors would like to thank all participants in the present study.

Ethical Considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Funding

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

Conflict of Interest

The authors declare no conflict of interest.

مقایسه تحلیل پدیدارشناسانه تجربیات معلمان درباره مشکلات دانش‌آموزان دوزبانه با بهره‌گیری از روش کلایزی و زنجیره تفکر GPT-4

آزاد الله کرمی استادیار، گروه آموزش علوم تربیتی، دانشگاه فرهنگیان، تهران، ایران. رایانامه: A.allahkarami@cfu.ac.ir

فریبا درتاج * نویسنده مسئول، گروه علوم تربیتی، دانشگاه پیام نور، تهران، ایران. رایانامه: F.D@pnu.ac.ir

چکیده

پژوهش حاضر باهدف بررسی مشکلات دانش‌آموزان دوزبانه و مقایسه‌ی تحلیل انسانی و ماشینی از تجارب معلمان پایه اول ابتدایی در این زمینه انجام شده است. رویکرد پژوهش، کیفی و روش آن پدیدارشناسی توصیفی کلایزی است. داده‌ها از طریق مصاحبه‌های نیمه‌ساختاریافته با ۱۵ معلم پایه اول مناطق دوزبانه اورامانات گردآوری و سپس با دو روش تحلیل شدند: (۱) تحلیل انسانی بر پایه‌ی مراحل هفت‌گانه کلایزی، و (۲) تحلیل ماشینی با بهره‌گیری از مدل GPT-4 و تکنیک زنجیره تفکر. در تحلیل انسانی، چهار مقوله اصلی شامل مشکلات آموزشی، زبانی، روان‌شناختی و اجتماعی شناسایی شد. در تحلیل ماشینی، داده‌ها در قالب شش مضمون طبقه‌بندی شدند: مشکلات زبانی پایه‌ای، تجربه‌های روانی منفی، شکاف فرهنگی، ناهمخوانی خانه و مدرسه، کناره‌گیری آموزشی، و فقدان حمایت نهادی. تحلیل تطبیقی نشان داد که اگرچه هم‌پوشانی قابل‌توجهی میان دو روش وجود دارد، اما تحلیل انسانی با درک عمیق‌تر زمینه فرهنگی و ادغام مفاهیم مرتبط، توانست تبیینی یکپارچه‌تر ارائه دهد. در مقابل، GPT-4 بادقت بالا در شناسایی الگوهای متنی و تم‌های مستقل، ابعاد نوینی از مسئله را آشکار ساخت. نتایج نشان داد در صورتی که خروجی‌های ماشینی توسط تحلیل‌گران انسانی بازبینی و با بافت نظری سنجیده شوند، ترکیب این دو می‌تواند مکمل مفیدی باشد؛ در غیر این صورت خطر برداشت‌های سطحی وجود دارد.

کلیدواژه‌ها: زبان مادری، مشکلات آموزشی، تحلیل پدیدارشناسی، GPT-4.

استناد به این مقاله: الله کرمی، آزاد، و درتاج، فریبا. (۱۴۰۵). مقایسه تحلیل پدیدارشناسانه تجربیات معلمان درباره مشکلات دانش‌آموزان دوزبانه با بهره‌گیری از روش کلایزی و زنجیره تفکر GPT-4. *فصلنامه اندازه گیری تربیتی*، *۱۷(۶۳)*، ۳۵-۱. ۱۰,۲۲۰۵٤/jem.2025.86707.3631

مقدمه

بر پایه گزارش‌های یونسکو، حدود نیمی از کودکان جهان آموزش رسمی خود را به زبانی غیر از زبان مادری آغاز می‌کنند؛ این در حالی است که بیش از ۷۰ درصد کشورهای جهان دارای ساختار زبانی متکثر و جوامعی چندزبانه‌اند. چنین ناهماهنگی زبانی، یکی از موانع اساسی در مسیر تحقق شعار جهانی «آموزش برای همه» تلقی می‌شود (الله کرمی، ۱۳۹۷). در چنین بستر چندزبانه‌ای، انتخاب زبان آموزش، به‌ویژه در سال‌های آغازین تحصیل، نقشی تعیین‌کننده در موفقیت یا ناکامی تحصیلی دانش‌آموزان ایفا می‌کند (Diba & Priyadarshini, 2023; Obiegbu, 2016). هر چند در کشورهایی مانند ایران، زبان مادری کودکان در برنامه‌های آموزشی نادیده گرفته می‌شود و آموزش به زبان غالب، به‌جای آموزش دوزبانه، زمینه‌ساز نابرابری‌های آموزشی شده است (سبحانی سنجبد، ۱۳۹۵)، در دنیا از دهه ۱۹۶۰ به بعد و با شکل‌گیری جنبش‌های حقوق مدنی، نگرش پژوهشگران نسبت به زبان‌آموزی تغییر یافته و مزایای زبانی، شناختی و اجتماعی زبان مادری و دوزبانگی مورد توجه بیشتری قرار گرفته است (Titone, 2013). زبان مادری نه صرفاً زبان پدر یا مادر، بلکه نخستین زبان آموخته‌شده در تعامل با محیط است (Alabi, 2002) و آموزش مبتنی بر آن می‌تواند به رشد شناختی، زبانی و هویتی کودک کمک کند. نیتو (Nieto, 2000) آموزش دوزبانه را بهره‌گیری همزمان و متعادل از دو زبان در موقعیت‌های آموزشی می‌داند، در حالی که بیکر (Baker, 2008) با نگاهی گسترده‌تر، دوزبانگی را پدیده‌ای پیچیده با ابعاد فردی، اجتماعی، فرهنگی و تحصیلی تلقی می‌کند و آن را در قالب‌هایی چون دوزبانگی هم‌زمان، افزایشی، کاهشی، فرهنگی و مدرسه‌ای طبقه‌بندی می‌نماید. سایر صاحب‌نظران مانند اسکات ناب‌کانگاس (Skutnabb-Kangas, 1981)، مالمبرگ (Malmberg, ۱۹۷۲)، نرسیسیانس (۱۳۸۲)، و بیلی‌استوک (Bialystok, 2007) نیز با بررسی انواع و پیامدهای دوزبانگی، نشان داده‌اند که چگونگی مواجهه نظام آموزشی با زبان مادری، تأثیر مستقیمی بر موفقیت تحصیلی و رشد روان‌شناختی کودکان دارد. در ایران، اغلب دانش‌آموزان دوزبانه نه دوزبانه طبیعی، بلکه دوزبانه مدرسه‌ای‌اند و معمولاً به‌جای تسلط کامل، در وضعیت «شبه دوزبانه همزاد» قرار دارند. این وضعیت در مناطق مرزی چون اورامانات، که زبان مادری دانش‌آموزان با زبان رسمی آموزش تفاوت دارد، بسیار حادتر است. تمرکز صرف بر زبان فارسی در نظام آموزش رسمی، در کنار حذف کامل زبان مادری

از برنامه درسی، موجب شده این دانش‌آموزان نه‌تنها از مزایای دوزبانگی محروم بمانند، بلکه دچار آسیب‌های آموزشی و روان‌شناختی متعددی شوند (الله کرمی، ۱۳۹۷).

درحالی‌که متخصصان معتقدند دوزبانگی عامل تقویت رشد شناختی، زبان‌شناختی و اجتماعی است، در بافت آموزشی ایران عمدتاً پیامدهای منفی آن موردتأکید قرار گرفته است؛ پیامدهایی که بیش از آنکه ناشی از خود دوزبانگی باشند، محصول فقدان سیاست‌های آموزشی حمایتی‌اند. ازجمله اختلال در درک مفاهیم پایه، افت تحصیلی (رحیم‌آبادی، ۱۳۸۳؛ واحدی، ۱۳۹۰)، مشکلات زبانی و شناختی (فیاضی و همکاران، ۱۳۹۳)، اضطراب و کناره‌گیری اجتماعی (کاظمی، ۱۳۸۶) و در نهایت ترک تحصیل و سرخوردگی تحصیلی (عصاره، ۱۳۸۷). این در حالی‌ست که پژوهش‌ها نشان داده‌اند دوزبانگی اگر به شکلی صحیح مدیریت شود، نه‌تنها مانع نیست بلکه می‌تواند موجب رشد توانش زبانی، تفکر واگرا، مهارت حل مسئله، و اعتمادبه‌نفس شود (Diaz, 1985; Gunning, 1981; Konaka, ۱۹۹۷; Albert et al, 2002; Obiegbu, 2016; Diba & Priyadarshini, 2023).

باوجود اهمیت موضوع، در بررسی پیشینه پژوهش در ایران می‌توان دریافت که مطالعات کیفی به‌ویژه با رویکرد پدیدارشناسی توصیفی درباره تجربه زیسته معلمان از مشکلات آموزشی دانش‌آموزان دوزبانه بسیار اندک است. اکثر پژوهش‌های موجود یا به روش کمی و در قالب آزمون‌های استاندارد انجام شده‌اند یا به تحلیل سیاست‌های کلان آموزشی پرداخته‌اند، درحالی‌که نگاه معلمان به‌عنوان ناظران و کنشگران اصلی فرایند یاددهی – یادگیری در مناطق دوزبانه، می‌تواند به تبیین بهتر و واقع‌گرایانه‌تر مسئله کمک کند.

در این میان، گسترش روزافزون هوش مصنوعی در حوزه‌های آموزشی و روان‌شناختی، ضرورت سنجش نگرش و سواد مرتبط با این فناوری را در جوامع گوناگون برجسته ساخته است؛ به‌گونه‌ای که پژوهش‌های اخیر در ایران به اعتباریابی ابزارهای اندازه‌گیری نگرش کلی به هوش مصنوعی (صدری دمیرچی و همکاران، ۱۴۰۴) و سواد هوش مصنوعی در میان دانشجویان (قدم‌پور و همکاران، ۱۴۰۴) پرداخته‌اند. با این حال، علی‌رغم این پیشرفت‌ها، کاربرد خودِ مدل‌های زبانی پیشرفته‌ای همچون GPT-4 در تحلیل داده‌های کیفی و درک عمیق تجاربِ انسانی، به‌ویژه در بافت‌های فرهنگی-آموزشی پیچیده‌ای مانند مناطق دوزبانه، همچنان به‌صورت نظام‌مند مورد بررسی قرار نگرفته است. این مسئله پرسش اساسی را مطرح می‌سازد که آیا تحلیل ماشینی داده‌های کیفی می‌تواند در شناسایی تجارب زیسته و استخراج ساختار پدیده‌های آموزشی نظیر مشکلات دانش‌آموزان دوزبانه مؤثر واقع شود؟

در سال‌های اخیر، توسعه مدل‌های زبانی بزرگ دسترسی به ابزارهای استخراج و خوشه‌بندی داده ها را افزایش داده است؛ با این وجود تاثیر عملی آن بر کیفیت تفسیرهای عمیق کیفی همچنان محل بحث است. به‌همین‌دلیل پژوهشگران از آن‌ها برای تسهیل گام‌های مقدماتی تحلیل کیفی بهره می‌گیرند (Bender et al., 2021; Birhane et al., 2022; Isoaho et al, 2021). پردازش زبان طبیعی که بر تعامل میان رایانه و انسان از طریق زبان طبیعی تمرکز دارد و هدف آن، توانمندسازی رایانه‌ها برای درک، تفسیر و تولید زبان انسانی به‌گونه‌ای معنادار است (Chowdhury, 2020)، در سال‌های اخیر دگرگونی قابل‌توجهی را تجربه کرده است؛ به‌گونه‌ای که از تمرکز اولیه بر تحلیل فراوانی واژگان به سمت درک معانی واژگان از طریق بردارهایی معنایی (Grimmer et al., 2022؛ Mikolov et al., ۲۰۱۳) پیش رفته است، بردارهای معنایی، واژه‌ها را به‌صورت بردارهای عددی در فضایی ریاضی ترسیم می‌کنند و روابط معنایی میان آن‌ها را باتوجه‌به زمینه و مجاورتشان بازنمایی می‌کنند؛ فرایندی که موجب افزایش توانایی ماشین‌ها در درک و پردازش زبان می‌شود.

تحولات اخیر در حوزه مدل‌های زبانی، به‌ویژه با معرفی معماری ترنسفورمر (Vaswani et al., 2017)، انقلابی در توانایی ماشین‌ها برای پردازش و تولید زبان طبیعی پدید آورده است. برخلاف مدل‌های سنتی که فقط به تعداد محدودی از واژگان پیشین توجه دارند و در درک وابستگی‌های بلندمدت ناتوان‌اند (Bengio et al., 2000)، مدل‌های زبانی بزرگ با استفاده از یادگیری عمیق و داده‌های متنی گسترده، ساختارهای معنایی پیچیده را فرا می‌گیرند. ترنسفورمرها با بهره‌گیری از سازوکار توجه، امکان تحلیل دقیق‌تر روابط معنایی در توالی‌های متنی را فراهم می‌سازند (Vaswani et al., 2017) و زمینه را برای تولید متونی منسجم و بافت‌محور (Devlin et al., 2018) فراهم کرده‌اند.

مدل GPT-3 که در ژوئن ۲۰۲۰ توسط OpenAI معرفی شد، نقطه عطفی در توسعه مدل‌های زبانی بزرگ به‌شمار می‌آید و نمونه‌ای برجسته از مقیاس‌پذیری معماری ترنسفورمر محسوب می‌شود که اساس مدل‌های پیشرفته‌تر بعدی را شکل داد (Brown et al., 2020). این مدل با ۱۷۵ میلیارد پارامتر طراحی شده و هدف آن پیش‌بینی واژه یا دنباله‌ای از واژگان بر پایه اطلاعات زمینه‌ای است. حجم بالای پارامترها، که روابط پیچیده درون داده‌ها را بازنمایی می‌کنند، در کنار آموزش گسترده بر مجموعه‌ای متنوع از منابع متنی، به GPT توانایی شگرفی در درک و تولید زبان طبیعی بخشیده است. برخی مطالعات موردی نشان

می‌دهند در وظایف خاص، مانند تولید متن ساختاریافته، عملکرد مدل‌های زبانی نردیک به عملکرد انسان بوده است (Binz & Schulz, 2023; Elkins & Chun, 2020). در عین حال، برخی پژوهش‌ها نشان داده‌اند که تلاش برای شباهت زباد خروجی مدل به زبان انسانی ممکن است در برخی وظایف مانند جزئیات مهم یا دقت علمی موجب افت عملکرد شود (Chen et al., 2023; Ye et al., 2023). پیشینه پژوهش‌ها همچنین بر توانایی چندوظیفه‌ای این مدل‌ها بدون نیاز به تنظیم خاص تأکید دارد (Radford et al., 2019). با این حال، دستیابی به عملکرد دقیق در هر وظیفه، مستلزم طراحی دقیق پرامپت‌هاست. در این زمینه، استراتژی‌هایی نظیر پرامپت‌های چندمثالی (Brown et al., 2020) و زنجیره تفکر که فرایند استدلال مرحله‌به‌مرحله را تسهیل می‌کنند (Wei et al., 2022)، مورد توجه قرار گرفته‌اند. هر چند نرم‌افزارهای تحلیل داده‌های کیفی نظیر Nvivo و Dedoose از اواخر دهه ۱۹۸۰ وارد عرصه پژوهش‌های کیفی شده‌اند (Creswell & Poth, 2018)، این ابزارها عمدتاً پشتیبان پژوهشگر در جنبه‌های فنی، سازمان‌دهی و مدیریت داده‌ها هستند، در حالی که بخش‌های پیچیده‌تر فرایند تحلیل، مانند تفسیر معنایی و استنتاج مفهومی، همچنان بر عهده تحلیل‌گران انسانی باقی مانده است (Marshall & Rossman, 2015). در مقابل، مدل‌های زبانی بزرگ به دلیل توانمندی چشمگیر خود در درک و تولید زبان طبیعی، پتانسیل آن را دارند که در بخش‌هایی از تحلیل کیفی، نظیر شناسایی تم‌ها یا ارائه تفسیرهای اولیه، به کار گرفته شوند. با این وجود، همچنان پرسش‌های بنیادینی در خصوص دقت، انسجام معنایی، و توانایی درک زمینه‌ای این مدل‌ها نسبت به تحلیل انسانی وجود دارد (Weidinger et al., 2021). افزون بر این، از آنجا که مدل‌های زبانی بزرگ عمدتاً با استفاده از متون تولیدشده توسط انسان آموزش می‌بینند، به‌طور ناگزیر بازتاب‌دهنده سوگیری‌هایی هستند که در بسترهای اجتماعی، تاریخی و فرهنگی بشر ریشه دارند؛ از جمله تعصبات نژادی، جنسیتی و مذهبی (Bhardwaj et al., 2020; Brown et al., 2020). برای مثال، در یک مطالعه تجربی که طی آن از GPT-3 خواسته شد گروه‌های نژادی را توصیف کند، مشخص شد که مدل، گروه سیاه‌پوستان را با واژگانی منفی‌تر نسبت به سایر گروه‌های نژادی توصیف کرده است (Brown et al., 2020). شواهد مشابهی درباره بازتولید کلیشه‌های جنسیتی و مذهبی نیز گزارش شده است (Floridi & Chiriatti, ۲۰۲۰).

اگرچه قابلیت مدل‌های مبتنی بر GPT در زمینه تحلیل داده‌های کمّی مثل به‌کارگیری مفسر کد[1](OpenAI, 2023) به طور تجربی بررسی شده و عملکرد آن‌ها با تحلیل‌گران تازه‌کار قابل‌مقایسه ارزیابی شده است، بااین‌حال، بهره‌گیری از این مدل‌ها در تحلیل داده‌های کیفی هنوز در مراحل ابتدایی پژوهشی قرار دارد و به‌ندرت به طور سیستماتیک بررسی شده است. در سال‌های اخیر، گام‌هایی در جهت تلفیق فناوری‌های پردازش زبان طبیعی با روش‌های تحلیل کیفی برداشته شده است، به‌عنوان نمونه، Gamieldien و همکاران (۲۰۲۳) از روش خوشه‌بندی مبتنی بر NLP برای گروه‌بندی داده‌های متنی بر اساس شباهت‌های معنایی استفاده کردند و سپس مدل GPT-3.5 را برای برچسب‌گذاری تماتیک خوشه‌ها به کار گرفتند. در مطالعه‌ای دیگر، Hamilton و همکاران (۲۰۲۳) ابتدا نقل‌قول‌های کلیدی را به‌صورت دستی استخراج کرده و پس از آن از ChatGPT خواستند تا تم‌های اصلی موجود در داده‌ها را شناسایی کند. این فرایندهای پیش‌پردازش، هرچند در جهت تسهیل تحلیل هستند، اما ممکن است از پیچیدگی تحلیلی بکاهند و بر دقت یا عمق تفسیر نهایی تأثیر بگذارند. علاوه بر این، مطالعه‌ای توسط Siiman و همکاران (۲۰۲۳) کاربرد ChatGPT را در کدگذاری استقرایی و قیاسی بررسی کرد و نشان داد که استفاده از این مدل در قالب تعامل پرسش‌وپاسخ می‌تواند فرایند تحلیل را تسهیل کند. تمامی این مطالعات به ارزیابی مقایسه‌ای عملکرد مدل‌های زبانی در کنار تحلیل‌گران انسانی پرداخته‌اند و در مجموع بر پتانسیل LLM ها در تحلیل داده‌های کیفی تأکید کرده‌اند.

در بسیاری از پژوهش‌های پیشین، تمرکز اصلی بر این بوده است که سوگیری‌های درونی مدل‌های زبانی چگونه می‌توانند بر تفسیر داده‌های کیفی تأثیر بگذارند و این مدل‌ها تا چه اندازه توانایی بازنمایی دقیق تجربیات انسانی را در بافت‌های خاص و متفاوت دارند. این پرسش به‌ویژه زمانی اهمیت می‌یابد که پای تحلیل‌های پدیدارشناسانه در میان باشد؛ تحلیلی که بر مفاهیمی چون «تجربه زیسته» و «قصدیت» تأکید دارد و مستلزم حفظ پیوند عمیق با زمینه‌های معنایی و زیستی مشارکت‌کنندگان است. از سوی دیگر باید پذیرفت فناوری‌های پردازش زبان طبیعی تنها زمانی می‌توانند در تحلیل داده‌های کیفی سودمند باشند که به‌عنوان «کمک‌یار تحلیلی» و نه جایگزین تحلیل‌گر انسانی تلقی شوند. بر این اساس، مقایسه و تلفیق دو مسیر تحلیلی (یکی بر پایه شهود و درک انسانی و دیگری متکی بر الگوریتم‌های زبانی) نه تنها به شفاف‌تر شدن مرزها و ظرفیت‌های هر رویکرد در مواجهه با داده‌های پیچیده کمک

---

[1]. code interpreter

می‌کند، بلکه امکان بررسی نحوه هم‌افزایی و تکمیل یکدیگر را در دستیابی به فهم عمیق‌تر واقعیت‌های آموزشی، به‌ویژه در حوزه‌هایی مانند مسائل زبانی و فرهنگی دانش‌آموزان دوزبانه، فراهم می‌سازد. برای نیل به این هدف باید پذیرفت تفسیر نهایی همچنان توسط پژوهشگران انجام شود. این رویکرد باعث می شود نتایج دو روش در کنار هم قرار گیرند، شباهت‌ها و تفاوت‌ها آشکار شوند و ظرفیت‌ها یا کاستی‌های رویکرد ماشینی نسبت به تحلیل پدیدارشناسانه توسط پژوهشگران سنجیده شود. این مقایسه‌ها می‌توانند نشان دهند که مدل‌های ماشینی تا چه اندازه قادرند درک عمیق‌تری از داده‌های کیفی ارائه دهند، یا در چه شرایطی می‌توان آن‌ها را به‌عنوان مکمل، و نه جایگزین، تحلیلگران انسانی به‌کار گرفت. همان‌گونه که براون و کلارک (Braun & Clarke, 2006) نیز تأکید کرده‌اند، تحلیل کیفی فرایندی تفسیری، بافت‌محور و پیچیده است که همواره به داوری انسانی نیاز دارد؛ ازاین‌رو، استفاده از ابزارهای هوش مصنوعی در این حوزه علاوه بر ارزیابی دقت، مستلزم تأمل اخلاقی و معرفت‌شناختی نیز هست.

مطالعه حاضر با تکیه بر همین دغدغه‌ها در پی آن است که با مقایسه‌ای نظام‌مند میان تحلیل انسانی و تحلیل ماشینی، به ارزیابی علمی ظرفیت مدل‌های زبانی بزرگ در پژوهش‌های کیفی بپردازد. بدین منظور، پژوهش حاضر در مرحله اول با اتخاذ رویکرد پدیدارشناسی توصیفی و روش کلایزی، به بررسی تجربیات معلمان درباره مشکلات آموزشی دانش‌آموزان دوزبانه در منطقه اورامانات پرداخت. در مرحله بعد، تحلیل‌های انسانی با تحلیل‌های به‌دست‌آمده از GPT-4 مقایسه گردید تا روشن شود این مدل تا چه حد می‌تواند در بازنمایی تجربه انسانی در سطح داده‌های کیفی، مکمل یا متمایز از تفسیر پدیدارشناسانه عمل کند. بنابراین، هدف اصلی این پژوهش درک پدیدارشناسانه تجربیات معلمان و هم‌زمان مقایسه تحلیل انسانی بر پایه روش کلایزی با تحلیل ماشینی مبتنی بر GPT-4 است.

روش

این پژوهش با رویکرد کیفی و به روش پدیدارشناسی توصیفی انجام شده است تا درک عمیق‌تری از تجربیات زیسته معلمان دانش‌آموزان دوزبانه پایه اول در منطقه اورامانات فراهم شود. هدف اصلی این روش بررسی چندلایه ادراک، نگرش و احساسات افراد نسبت به مشکلات آموزشی این گروه در یک محیط طبیعی و بدون پیش‌فرض بود. روش پدیدارشناسی انتخاب شد؛ چراکه امکان استخراج لایه‌های پنهان تجربه را فراهم می‌آورد و

با تمرکز بر معنای عینی زندگی شرکت‌کنندگان، چارچوب مناسبی برای مطالعه فراهم می‌سازد. در این پژوهش، داده‌ها از طریق مصاحبه‌های نیمه‌ساختاریافته با معلمان دانش‌آموزان دوزبانه پایه اول طی سال تحصیلی ۱۴۰۴-۱۴۰۳ جمع‌آوری شد. شرکت‌کنندگان به‌صورت هدفمند انتخاب شدند. فرایند گردآوری داده‌ها تا رسیدن به اشباع نظری ادامه یافت؛ به‌طوری که اشباع در مصاحبه‌ی سیزدهم حاصل شد. با این حال، برای تقویت قابلیت اطمینان یافته‌ها، تا مصاحبه پانزدهم ادامه یافت. راهنمای مصاحبه با توجه به اهداف پژوهش تدوین و پس از انجام سه مصاحبه آزمایشی، اصلاح گردید. مصاحبه‌ها به‌طور انفرادی و با مدت زمانی بین ۳۰ تا ۴۰ دقیقه انجام شدند، ضبط صوتی کاملاً با رضایت شرکت‌کنندگان انجام و متن‌های پیاده‌سازی‌شده تحلیل شدند. سؤالات نهایی مصاحبه عبارت بودند از:

۱. مهم‌ترین مشکلات دانش‌آموزان دوزبانه در فرایند یادگیری کدام‌اند؟
۲. تفاوت زبان مادری و زبان آموزش رسمی چه تأثیری بر عملکرد تحصیلی و اعتمادبه‌نفس دارد؟
۳. تعاملات در مدرسه در بستری که زبان مادری با زبان رسمی متفاوت است چه اثری بر یادگیری و روابط اجتماعی دانش‌آموزان دارد؟
۴. چه حمایت‌هایی از طرف خانواده، مدرسه یا نظام آموزشی می‌تواند برای این دانش‌آموزان مؤثر باشد؟
۵. آیا محتوای آموزشی، شیوه تدریس یا ارزیابی‌ها با نیازهای زبانی و فرهنگی این دانش‌آموزان هماهنگ است؟ در صورت منفی بودن، چه تغییراتی پیشنهاد می‌کنید؟

به‌تمامی شرکت‌کنندگان کد اختصاص داده شد تا گمنامی حفظ شود و زمان و مکان گفتگو مطابق ترجیح آنان تنظیم گردد. تحلیل داده‌ها در دو مسیر موازی اجرا شد. در مسیر اول، تحلیل انسانی با هفت مرحله چرخه کلایزی انجام شد: (۱) درک کلی داده با خوانش مکرر متن، (۲) استخراج گزاره‌های معنادار، (۳) فرموله‌سازی معانی، (۴) خوشه‌بندی معانی به تم‌های مفهومی، (۵) نگارش شرح جامع پدیده، (۶) استخراج ساختار بنیادین تجربه، (۷)

بازبینی و تأیید نهایی یافته‌ها با مشـــارکت‌کنندگان. برای ارتقای اعتبار و قابلیت اعتماد، پژوهشـگر پیش‌فرض‌های خود را قبل از تحلیل ثبت و با روش تعلیق فرض‌ها کنترل کرده، ســپس در مراحل تحلیل از آن‌ها به طور مداوم بازنگری نمود. نتایج اولیه برای بازخورد مشارکت‌کنندگان ارسال و اصلاحات حاصل شد. یافته‌ها نیز توسط دو متخصص مستقل در حوزه روش‌شناسی کیفی و زبان‌شناسی بررسی و تأیید گردید. در مسیر دوم، از مدل زبانی GPT-4 از طریق محیط وب ChatGPT (OpenAI) و روش زنجیره تفکر بدون نمونه[1] اســتفاده شــد. پس از پیاده‌ســازی مصــاحبه‌ها، متن آن‌ها در قالب یک فایل ورد ذخیره و بارگذاری گردید. پیش از بارگذاری، نام‌ها و مشــخصــات فردی خذف شــده و املای جملات یکسان‌سازی گردید.کلیه مصاحبه‌ها به‌صورت متنِ اصلی (فارسی) و بدون ترجمه پیشین وارد مدل شد؛ هرگونه کدسوئیچینگ و اصطلاحات بومی نیز حفظ گردید تا معانی اصلی از بین نرود. در این پژوهش تحلیل ماشینی در شانزدهم خرداد ۱۴۰۴ انجام شد. لازم به ذکر اسـت نقش این مدل در پژوهش حاضـر صـرفاً به‌عنوان ابزار کمکی برای اسـتخراج گزاره‌های معنادار و پیشــنهاد تم‌های اولیه تعریف گردید و تفســیر نهایی همواره توســط پژوهشگران انجام خواهد شد. از آنجا که در نسخه وب ChatGPT پژوهشگر امکان تغییر پارامترهای فنی (مانند temperature، top_p، max_tokens، و غیره) را ندارد، کلیه تحلیل‌ها بر اســاس تنظیمات پیش‌فرض پلتفرم OpenAI انجام گرفت. این موضــوع در بخش روش پژوهش به‌طور شـفاف گزارش می‌شـود تا امکان بازتولیدپذیری مطالعه فراهم گردد. برای شفافیت بیشتر، پرامپت اصلی به‌کاررفته در پژوهش عیناً در زیر ارائه می‌شود:

«نقش شـما تحلیل‌گر کیفی تربیتی با تخصـص در روش پدیدارشـناسـی توصـیفی کلایزی اسـت. متن آپلود شـده برگرفته از مصـاحبه‌هایی اسـت که با معلمان دانش‌آموزان پایه اول ابتدایی به‌منظور شناسایی مشکلات آموزشی دانش‌آموزان دوزبانه انجام شده است. لطفاً تحلیل را در ۷ مرحله‌ی زیر انجام دهید: (۱) ابتدا متن را به‌دقت بخوانید و بادقت آشــنا

[1]. Zero-Shot Chain-of-Thought prompting

شوید. (۲) جملات یا عبارات کلیدی که بیانگر مشکل هستند را شناسایی کنید، پژوهش در پی پاسخگویی به این پرسش است: مهم‌ترین مشکلات دانش‌آموزان دوزبانه پایه اول ابتدایی کدام‌اند؟ (۳) برای هر گزاره معنادار، معنی ضمنی آن را در قالب یک عبارت کوتاه ارائه دهید. (۴) معانی مشابه را جمع‌آوری و آن‌ها را در چند تم مفهومی گروه‌بندی کنید. (۵) برای هر تم، یک توصیف جامع از مشکل موردبررسی بنویسید. (۶) ساختار کلی مشکلات دانش‌آموزان دوزبانه را در قالب چند جمله بیاورید و توصیف جامعی از تجربه زیسته‌ی مشارکت‌کنندگان ارائه دهید. (۷) برای هر تم، یک یا دو نقل‌قول کامل از متن مصاحبه ارائه دهید تا تم را مستند کند». نهایتاً نوشته شد: (Let's think step by step). این عبارت برای فعال‌سازی استدلال گام‌به‌گام گنجانده شد.

برای آزمون حساسیتِ پرامپت و ثبات نتایج سه نسخه متفاوت از پرامپت اجرا شد: (۱) پرامپت اصلی، (۲) پرامپت بافت‌محور که در آن به‌طور صریح از مدل خواسته شد به اصطلاحات محلی و کدسوئیچینگ توجه کند و (۳) پرامپت نمونه‌دار[۱] که همراه با دو نمونه راهنما ارائه شد. در پرامپتِ بافت‌محور صریحاً از مدل خواسته شد هنگام تحلیل به اصطلاحات بومی، کدسوئیچینگ و معانی محلی توجه کند و در مواردی که چند برداشت ممکن است وجود داشته باشد، توضیح کوتاه ارائه دهد. به این ترتیب، هم از روش نمونه‌دار برای هدایت مدل استفاده شد و هم با دستور مستقیم حساسیت به زمینه زبانی و فرهنگی در مدل تقویت گردید. همچنین در پرامپتِ نمونه‌دار دو مثال راهنما به مدل داده شد. هر مثال شامل یک گزیده کوتاه از متن مصاحبه به‌همراه تمی بود که توسط پژوهشگر انسانی برای آن گزیده تعیین شده بود. هدف از این کار، آشناکردن مدل با شکل و سطح تحلیل مورد انتظار و کاهش احتمال برداشت نادرست از عبارات محلی یا مبهم بود. متن مصاحبه ها با هر سه پرامپت اجرا شد و خروجی خام هر سه حالت برای ارزیابی میزان همپوشانی مفهومی

---

۱. few-shot

مورد بررسی قرار گرفت. مقایسه نشان داد که تم‌های اصلی در هر سه نسخه مشترک بودند. این نتایج بیانگر ثبات نسبی خروجی‌های GPT-4 در شرایط اجرای پیش‌فرض است.

برای مقایسه نتایج حاصل از تحلیل پدیدارشناسانه انسانی و تحلیل ماشینی مبتنی بر GPT-۴، از طرح هم‌زمانِ موازی استفاده شد؛ بدین معنا که هر مسیر تحلیلی به‌صورت مستقل اجرا شد و سپس با بهره‌گیری از یک چارچوب مقایسه تم‌ها کنار هم قرار گرفتند. این چارچوب شامل شاخص‌هایی چون سطح تحلیل، مبنای گروه‌بندی، حساسیت زمینه‌ای، چارچوب نظری و انعطاف‌پذیری مفهومی بود. هدف از این مقایسه، سنجش توانایی GPT-4 در تحلیل داده‌ها بود؛ اما نه به قصد جایگزین کردن تحلیل انسانی، بلکه برای تکمیل و مقایسه آن.

یافته‌ها

در ادامه تحلیل داده‌های کیفی حاصل از مصاحبه‌های نیمه‌ساختارمند با معلمان در چهار بخش ارائه خواهد شد. (۱) نتایج تحلیل انسانی، (۲) تحلیل ماشینی، (۳) مقایسه و تحیل نقادانه.

الف) تحلیل انسانی: نتایج نشان داد: مشکلات دانش‌آموزان دوزبانه پایه اول مقطع ابتدایی به چهار دسته اصلی آموزشی، روان‌شناختی، اجتماعی و زبانی تقسیم می‌شود که در ادامه هرکدام از این مشکلات به همراه جملاتی از تحلیل محتوای کیفی دررابطه‌با آن‌ها توضیح داده خواهد شد.

آموزشی: مشارکت‌کنندگان تأکید داشتند که زبان آموزش رسمی، در بسیاری از موارد، با زبان فهم‌پذیر برای دانش‌آموزان دوزبانه فاصله دارد و این امر سبب ضعف در درک مفاهیم درسی، ناتوانی در اجرای تکالیف و کاهش مشارکت آموزشی می‌شود. یادگیری اولیه مفاهیم به زبان مادری، به‌ویژه در حوزه‌های شناختی پایه، نقشی اساسی در موفقیت تحصیلی این دانش‌آموزان دارد. تحلیل مصاحبه‌ها نشان داد در مواردی که واژگان یا مفاهیم دارای هم‌پوشانی معنایی میان زبان مادری و زبان رسمی بودند، دانش‌آموزان عملکرد بهتری داشتند. اما در مواجهه با مفاهیم ناآشنا یا فاقد معادل در زبان مادری، دچار سردرگمی و عقب‌ماندگی می‌شدند. یکی از مشارکت‌کنندگان اظهار داشت که در آغاز، زمانی که تنها

به زبان فارسی صحبت می‌کرد، اکثر دانش‌آموزان قادر به انجام تکلیف نبودند. اما به‌محض این‌که همان پرسش را به زبان کردی مطرح کرد، تقریباً همه‌ی دانش‌آموزان آن را درک کرده و به‌درستی اجرا نمودند. این تجربه نشان‌دهنده‌ی نقش کلیدی زبان مادری در فهم و یادگیری مفاهیم آموزشی در کلاس‌های چندزبانه است. دانش‌آموزان با کلماتی مثل "دەس" یا "ئاو"[1] بهتر ارتباط می‌گرفتند. نتایج نشان داد که نبود پیوند میان محتوای آموزشی و زبان مادری باعث اتلاف انرژی شناختی، افزایش اضطراب یادگیری و محدودشدن فرصت‌های واقعی برای درک عمیق مطالب می‌شود.

روان‌شناختی: استفاده‌نشدن از زبان مادری در مدرسه، علاوه بر اثرات آموزشی، پیامدهای روان‌شناختی مهمی برای دانش‌آموزان دوزبانه به همراه دارد. مشارکت‌کنندگان به‌ویژه بر تضعیف اعتمادبه‌نفس، بروز احساس طردشدگی، و فقدان امنیت روانی در محیط مدرسه تأکید داشتند. تحلیل داده‌ها نشان داد که فاصله زبانی میان خانه و مدرسه، منجر به گسست عاطفی دانش‌آموزان و کاهش انگیزش درونی آنان برای یادگیری می‌شود. یکی از مشارکت‌کنندگان در این باره می‌گوید: "وقتی درس رو فارسی می‌گفتم، بچه‌ها هی سرشون رو می‌بردن طرف هم با هم پچ‌پچ می‌کردن، می‌پرسیدن که "آقا معلم چی می‌گه؟" ولی تا همون جمله رو به زبان کُردی گفتم، یهو کلاس زنده شد، همه شروع کردن به خندیدن". دانش‌آموزان در این شرایط اغلب احساس ناتوانی، تحقیر و انزوای روانی را تجربه می‌کنند. تداوم این وضعیت می‌تواند به کاهش مشارکت در فعالیت‌های کلاسی، اجتناب از تعامل با معلم و در نهایت، فرسایش سرمایه روانی دانش‌آموزان منجر شود. بعلاوه استفاده از زبان مادری در آموزش، با تقویت پیوستگی زبانی و عاطفی، موجب بهبود کیفیت ارتباطات کلاسی و دسترسی بهتر دانش‌آموزان به ابزار اصلی تفکر یعنی زبان می‌شود. این دانش‌آموزان ممکن است به دلیل عدم‌پذیرش زبان و هویت فرهنگی خود احساس انزوا، اضطراب و کاهش اعتمادبه‌نفس کنند. محدودیت در توانایی برقراری ارتباط مؤثر با معلمان و هم‌کلاسی‌ها، فشار برای یادگیری زبان دوم در کنار زبان مادری و تجربه عدم موفقیت در تحصیل، می‌تواند زمینه‌ساز بروز استرس و مشکلات روانی شود. همچنین، فقدان حمایت‌های عاطفی و آموزشی مناسب، رشد عاطفی و اجتماعی آن‌ها را مختل می‌کند و در نهایت بر عملکرد تحصیلی و سلامت روانی این گروه تأثیر منفی می‌گذارد.

---

[1] . دەس در زبان کردی به معنای دست و ئاو به معنای آب است.

اجتماعی: یکی دیگر از تم‌های برجسته، موانع اجتماعی ناشی از عدم انطباق زبان آموزش با زبان مادری بود. زبان در این زمینه، نه به‌عنوان ابزار ارتباط، بلکه به‌عنوان مانعی برای تعامل اجتماعی در کلاس و مدرسه عمل می‌کرد. مشارکت‌کنندگان به مواردی اشاره کردند که زبان رسمی موجب انزوای اجتماعی، قطع ارتباطات همسالان، و کاهش تعاملات کلاسی شده بود. مشارکت‌کننده‌ای می‌گوید: "وقتی در شروع کلاس به زبان رسمی سلام و احوالپرسی کرده و خودم را معرفی کردم دانش‌آموزان فقط نگاه می‌کردند و زبان بدن آنها نشان می‌داد یه جای کار می لنگه". یافته‌ها نشان می‌دهد که نظام آموزشی هنگامی که زبان مادری و فرهنگ بومی دانش‌آموزان اقلیت زبانی را نادیده می‌گیرد یا به حاشیه می‌راند، به طور غیرمستقیم و ناخواسته زمینه‌ساز ایجاد «فقر ارتباطی» در میان این دانش‌آموزان می‌شود. این فقر ارتباطی به معنای کاهش فرصت‌ها و کیفیت تعاملات کلامی و غیرکلامی با معلمان، هم‌کلاسی‌ها و محیط مدرسه است که در نهایت موجب ضعف در برقراری ارتباط مؤثر و مشارکت فعال آن‌ها در فرایند یادگیری و زندگی مدرسه می‌گردد. از سوی دیگر، حذف یا کم‌اهمیت شمردن زبان مادری به منزله رد هویت فرهنگی و زبانی دانش‌آموزان است که می‌تواند موجب بروز احساس طردشدگی اجتماعی، کاهش عزت‌نفس و ضعف در شکل‌گیری هویت فردی و گروهی آنان شود. این مسائل، علاوه بر تأثیر منفی بر روند تحصیلی، می‌تواند پیامدهای روان‌شناختی همچون اضطراب، افسردگی و کاهش احساس تعلق به مدرسه و جامعه را به همراه داشته باشد. در نتیجه، عدم توجه به زبان و فرهنگ بومی دانش‌آموزان اقلیت، نه‌تنها بر توانایی‌های ارتباطی آنان اثر می‌گذارد، بلکه سلامت روانی و رشد اجتماعی آنان را نیز به‌طورجدی تهدید می‌کند.

زبانی: یادگیری زبان رسمی در شرایطی که زبان مادری در آموزش نادیده گرفته می‌شود، با خطاهای زبانی متعدد همراه است. مشارکت‌کنندگان به انواع خطاهای زبانی از جمله حذف، گزینش و ترتیب نادرست واژگان اشاره داشتند که ناشی از انتقال ساختارهای زبانی مادری به زبان دوم است. اگرچه این خطاها بخشی طبیعی از فرایند یادگیری زبان دوم هستند، اما معلمان و دانش‌آموزان دیگر اغلب آن‌ها را به دیده تمسخر یا ناتوانی نگاه می‌کنند که خود به تقویت احساس شرم و بازداری در یادگیری می‌انجامد. معلمی در مصاحبه‌ها می‌گوید: "یکی از بچه‌ها می‌خواست از من تعریف کنه و گفت: "آقا تو چقد خاسی[1]! " و متأسفانه بچه‌های دیگر مسخرش کردند". این در حالی است که کد

[1]. شما چقدر خوبید.

سوئیچینگ یعنی جابه‌جایی طبیعی بین دو زبان در حین صحبت که در کلاس‌های دوزبانه طبیعی است و خطاهای زبانی، صورت‌های زبانی نامطلوب نیستند، بلکه نشانگر طرز کار درونی فرایند یادگیری زبان است. معلم دیگری به تنوع گویش‌ها در زبان کُردی اشاره می‌کند و می‌گوید: "دانش‌آموزی هورامی داشتم که نه زبان رسمی را متوجه می‌شد و نه زبان عموم هم‌کلاسی‌ها". در این مورد فقدان آموزش دوزبانه و بی‌توجهی به تنوع زبانی در کلاس، منجر به طرد آموزشی و حاشیه‌نشینی زبانی این دانش‌آموزان می‌شود. این یافته‌ها نشان می‌دهد که برخورد نامناسب با خطاهای زبانی، موجب تضعیف جرئت ارتباطی و افزایش اضطراب یادگیری در دانش‌آموزان دوزبانه می شود. نبود آموزش معلمان در زمینه تفاوت های زبانی و نبود پذیرش زبان مادری به‌عنوان منبع مشروع یادگیری، از جمله چالش‌های جدی در آموزش زبان دوم برای این گروه از دانش‌آموزان است.

ب) تحلیل ماشینی: در مسیر دوم تحلیل داده‌ها، از توان محاسباتی مدل GPT-4 با استفاده از الگوریتم زنجیره تفکر بدون نمونه بهره گرفته شد تا داده های کیفی حاصل از مصاحبه‌های معلمان، از منظر رویکرد پدیدارشناسی توصیفی کلایزی تحلیل شود. براین‌اساس، هفت مرحله تحلیل به‌صورت ساختاریافته بر روی متن کامل مصاحبه‌ها اعمال گردید. در این فرایند، معانی ضمنی از گزاره‌های کدگذاری‌شدن استخراج و در قالب شش تم مفهومی اصلی طبقه‌بندی شدند.[1] این تم‌ها نشان‌دهنده چالش‌های اساسی دانش‌آموزان دوزبانه پایه اول ابتدایی در منطقه اورامانات هستند.[2]

گام ۱ و ۲: آشنایی با داده‌ها و استخراج گزاره‌های معنادار: در مرحله نخست، مدل با خوانش مکرر و دقیق متن مصاحبه ها، درکی کلی از فضای روانی، زبانی و فرهنگی دانش‌آموزان دوزبانه کسب کرد. سپس ۱۳۰ گزاره معنادار شناسایی شد که بیانگر تجربه زیسته معلمان از چالش‌های یادگیری این کودکان بود؛

- دانش‌آموزان می‌ترسند اگر اشتباه کنند یا سؤال بپرسند، دیگران آنها را مسخره کنند و کمتر در کلاس مشارکت می‌کنند.

---

۱ . همان‌گونه که در بخش روش‌شناسی اشاره شد، در این پژوهش از سه پرامپت متفاوت استفاده گردید. با توجه به شباهت و نزدیکی نتایج حاصل از این سه پرامپت، در این بخش صرفاً خروجی پرامپت اصلی گزارش می‌شود.

۲ . این بخش خروجی GPT-4 براساس پرامپت نوشته شده در بخش روش‌شناسی است.

- دانش‌آموزان احساس می‌کنند فرهنگ و زبان مادری‌شان در مدرسه جدی گرفته نمی‌شود و این باعث بی‌ارزشی آنها می‌شود.
- به دلیل ترس از قضاوت و تمسخر، شجاعت دانش‌آموزان برای مشارکت در جمع کاهش‌یافته است.

گام ۳: تبدیل گزاره‌ها به معانی ضمنی: هر گزاره‌ی کلیدی به یک معنای ضمنی روشن تبدیل شد. برای نمونه: گزاره کلیدی "کودک دوزبانه به دلیل ترس از اشتباه ایده خود را بیان نمی کند" معنای ضمنی "ترس از اشتباه زبانی" پیدا کرد و "ساختار ذهنی کودک متکی به زبان مادری و ناسازگار با فارسی رسمی است" به عنوان "گسست مفهومی-زبانی" مشخص شد. "کودک احساس می‌کند جایگاهی در فرهنگ مدرسه ندارد" به عنوان "احساس بی‌هویتی" و "نظام آموزشی ما برای زبان غالب طراحی شده و دیگر زبان‌ها را نادیده می‌گیرد" به عنوان "بی‌تفاوتی مدرسه به تنوع زبانی" نامگذاری شد.
گام ۴: گروه‌بندی معانی مشابه در قالب تم‌های مفهومی: معانی ضمنی به صورت استقرایی در ۶ تم اصلی زیر دسته‌بندی شدند:

۱. مشکلات زبانی پایه‌ای در درک و بیان زبان رسمی
۲. تجربه‌های هیجانی و روانی منفی ناشی از فشار زبانی و فرهنگی
۳. شکاف فرهنگی و احساس بیگانگی با محتوای آموزشی و محیط مدرسه
۴. ناهمخوانی زبانی و تربیتی میان خانه و مدرسه
۵. کناره‌گیری آموزشی و کاهش مشارکت فعال در کلاس
۶. فقدان حمایت ساختاری و بی‌توجهی نهادی به تفاوت‌های زبانی

گام ۵: نگارش توصیف جامع برای هر تم: مشکلات زبانی پایه‌ای در درک و بیان زبان رسمی: معلمان گزارش دادند که دانش‌آموزان دوزبانه در جمله‌سازی، واژه‌یابی و درک مفاهیم درسی با مشکل مواجه‌اند. ساختار متفاوت زبان مادری، کاربرد نادرست زمان افعال و تلاش برای ترجمه هم‌زمان موجب اختلال در یادگیری زبان رسمی و دروس مدرسه می‌شود.

📌نقل‌قول: بچه‌ها نمی‌توانند خوب صحبت کنند؛ چون ساختار جمله‌ها در زبان مادری‌شان فرق می‌کند.

تجربه‌های هیجانی و روانی منفی ناشی از فشار زبانی و فرهنگی: دانش‌آموزان اغلب به دلیل اشتباهات زبانی دچار ترس، اضطراب و احساس شرمندگی می‌شوند. این حالت‌ها موجب کاهش اعتمادبه‌نفس، سکوت تحمیلی و در نهایت کناره‌گیری از مشارکت آموزشی می‌گردد.

📌نقل‌قول: وقتی اشتباه می‌کنند، می‌ترسند که بقیه بخندند یا معلم تذکر بدهد؛ برای همین سکوت را ترجیح می‌دهند.

شکاف فرهنگی و احساس بیگانگی با محتوای آموزشی و محیط مدرسه: نبود نشانه‌های فرهنگی و زبانی دانش‌آموزان در مدرسه، محتوای درسی و مراسم‌ها، حس طرد و بیگانگی را تقویت کرده و دانش‌آموز را از محیط آموزشی جدا می‌سازد.

📌نقل‌قول: در مدرسه هیچ نشانی از زبان و فرهنگ بچه‌های ما نیست. حس می‌کنند اینجا به آن‌ها تعلق ندارد.

ناهماهنگی زبانی و تربیتی میان خانه و مدرسه: تضاد زبان خانه با زبان آموزش رسمی، موجب سردرگمی کودک در درک مفاهیم درسی، ضعف در انتقال یادگیری و تضاد رفتاری در محیط آموزشی می‌شود.

📌نقل‌قول: دانش‌آموز وقتی وارد کلاس می‌شود، دائم در ذهنش در حال ترجمه از زبان مادری به فارسی است و همین تمرکز او را کم می‌کند.

کناره‌گیری آموزشی و کاهش مشارکت فعال در کلاس: برخی دانش‌آموزان به دلیل ترس از قضاوت، بی‌اعتمادی به توانایی زبانی و تجربه‌های منفی قبلی، مشارکت فعالی در کلاس ندارند و تدریجاً از فعالیت‌های گروهی و گفتاری کنار می‌کشند.

فقدان حمایت ساختاری و بی‌توجهی نهادی به تفاوت‌های زبانی: از نگاه معلمان، نظام آموزشی ایران به تفاوت‌های زبانی بی‌توجه است. نبود برنامه منسجم برای آموزش زبان دوم و یکسان‌سازی فرهنگی، زمینه‌ساز تشدید مشکلات آموزشی شده است.

📌نقل‌قول: نظام آموزشی ایران نسبت به تنوع زبانی دانش‌آموزان در استان کرمانشاه بی‌توجه است و همه آن‌ها را از یک زاویه می‌بیند.

گام ۶: ساختار کلی مشکلات و توصیف تجربه زیسته: مشارکت‌کنندگان مشکلات دانش‌آموزان دوزبانه را ترکیبی از موانع زبانی، روانی، فرهنگی، اجتماعی و ساختاری دانستند. ورود به مدرسه، برای بسیاری از این دانش‌آموزان به معنای مواجهه با زبان و فرهنگی بیگانه است که نه‌تنها فرایند یادگیری را مختل می‌کند، بلکه موجب آسیب روانی و فاصله‌گیری آن‌ها از محیط آموزش می‌شود. در غیاب درک و حمایت نهادی، این تجربه منجر به فرسودگی تحصیلی، انزوا و حتی ترک مسیر آموزشی می‌گردد.

گام ۷: استناد به نقل‌قول‌هایی از متن (کدها)

تم ۱: چالش‌های زبانی و یادگیری زبان دوم: درک نادرست واژگان انتزاعی، کاربرد نادرست زمان گذشته.

تم ۲: تجربه روانی و هیجانی منفی: ترس از مسخره‌شدن در جمع کلاس، فشار روانی هم‌زمان با فشار تحصیلی.

تم ۳: شکاف فرهنگی و حذف هویت بومی: حذف غیررسمی زبان مادری، تجربه حذف فرهنگی

تم ۴: ناهماهنگی بین خانه و مدرسه: تفاوت سبک تربیتی خانواده با انتظار مدرسه، ناهمگونی زبان خانه با زبان مدرسه.

تم ۵: کاهش مشارکت و تعامل آموزشی: کاهش مشارکت در فعالیت‌های کلاسی، تمایل به سکوت برای پرهیز از خطا.

تم ۶: فقدان حمایت نهادی و معلم‌محور: بی‌توجهی مدرسه به تنوع زبانی، نبود الگوی موفق از اقلیت زبانی.

ج) مقایسه و تحلیل نقادانه

جدول ۱. علل تفاوت در تعداد تم‌های شناسایی شده بین دو روش تحلیل

عامل	تحلیل انسانی	تحلیل ماشینی	مثال شاهد از داده‌ها
سطح تحلیل	کل‌نگر با تأکید بر وحدت تجربه زیسته	جزءنگر با تفکیک دقیق ابعاد	تحلیل انسانی به عدم حمایت ساختاری بی‌توجهی کرده و آن را در تم آموزشی دیده بود.

مبنای گروه‌بندی	ارتباط ذاتی پدیده‌ها بر اساس پدیدارشناسی	فراوانی و هم رخدادی واژگان در متن	تحلیل انسانی «اضطراب» و «کناره‌گیری» را به‌عنوان جنبه‌های یک پدیده تلفیق کرده بود.
حساسیت زمینه‌ای	تفسیر مبتنی بر بافت فرهنگی اورامانات	تحلیل مستقل از زمینه با تمرکز بر الگوهای زبانی	بی‌توجهی به گویش‌های مختلف زبان کُردی توسط ماشین
چارچوب نظری	توجه به نظریه‌های دوزبانه مثل کامینز	فاقد چارچوب نظری مشخص	طبقه‌بندی در تحلیل GPT-4 غالباً بر اساس ساختار جمله بود.
انعطاف‌پذیری مفهومی	ادغام مضامین با همپوشانی بالا (مثلاً زبان و فرهنگ)	حفظ تمایز بین مقولات	در تحلیل انسانی یکپارچه‌سازی نابرابری آموزشی در تم اجتماعی قرار گرفت، درحالی‌که GPT-4 تم مستقل برای بی‌عدالتی‌های نهادی ایجاد کرد.

تفاوت در تعداد تم‌های شناسایی شده (۴ تم در تحلیل انسانی در مقابل ۶ تم در GPT-4)

نا شی از تفاوت‌های پارادایمی در رویکردهای تحلیل ا ست. تحلیل ان سانی با پیروی از روش کلایزی، بر «وحدت معنایی» تجارب تأکید دارد و مضـــامین را در چارچوب‌های کلی‌تر گروه‌بندی می‌کند. در مقابل، GPT-4 با اتکا به الگوریتم‌های آماری، به تفکیک دقیق‌تر مفاهیم بر اســاس شــاخص‌های ســطح متن (مانند تکرار واژگان، ســاختارهای نحوی) می‌پردازد. به‌عنوان‌مثال، درحالی‌که تحلیل انســـانی «احســـاس طردشـــدگی» و «کاهش مشـــارکت» را به‌عنوان جلوه‌های مختلف یک پدیده (شـــکاف هویتی) تفســـیر می‌کند، GPT-4 این موارد را در دو تم مجزا «تجر به های هیجانی» و «کناره‌گیری آموزشـــی» طبقه‌بندی می‌نماید. این تفاوت به‌ویژه در مواجهه با پیچیدگی‌های فرهنگی مشـــهود اســـت؛ جایی که تحلیل انســـانی قادر به تشـــخیص ارتباط بین «زبان مادری» و «احســـاس تعلق» اســـت، اما GPT-4 این مفاهیم را صـــرفاً بر اســـاس همایندی واژگانی گروه‌بندی می‌کند.

جدول ۲. تطبیق تم‌ها

تم‌های استخراج‌شده توسط GPT-4 (٦ مورد)	نزدیک‌ترین تم معادل در تحلیل انسانی (٤ مورد)	توضیح تطابق/عدم تطابق

مشکلات زبانی پایه‌ای در درک و بیان زبان رسمی	مشکلات زبانی	هر دو مستقیماً به چالش‌های زبانی دانش‌آموزان اشاره دارند.
تجربه‌های هیجانی و روانی منفی	مشکلات روان‌شناختی	تطابق کامل – هر دو به مسائل عاطفی پرداخته‌اند.
شکاف فرهنگی و احساس بیگانگی	مشکلات اجتماعی	تطابق کامل – تمرکز بر هویت فرهنگی.
ناهمخوانی زبانی و تربیتی بین خانه و مدرسه	مشکلات آموزشی	تحلیل انسانی این موضوع را در بستر آموزشی بررسی کرده بود.
کناره‌گیری آموزشی و کاهش مشارکت	مشکلات روان‌شناختی	تحلیل انسانی این رفتارها را پیامد مسائل روانی می‌دانست.
فقدان حمایت ساختاری و نهادی	پوشش داده نشد	این تم مستقل در تحلیل انسانی وجود نداشت.

تحلیل مقایسه‌ای نشان داد که درحالی‌که هر دو روش به مضامین مشترکی پرداخته‌اند، تفاوت‌های معناداری در سطح تحلیل و طبقه‌بندی وجود دارد. تم مشکلات زبانی پایه‌ای در **GPT-4** به طور خاص بر چالش‌های ساختاری زبان مانند خطاهای دستوری، دشواری در درک واژگان انتزاعی و انتقال ساختارهای زبانی از زبان مادری تمرکز داشت، درحالی‌که تحلیل انسانی این موضوع را در چارچوب گسترده‌تر مشکلات زبانی بررسی کرده بود که شامل تأثیر این چالش‌ها بر فرایند یادگیری و تعاملات کلاسی نیز می شد. تم تجربه‌های هیجانی منفی در GPT-4 و چالش‌های روان‌شناختی در تحلیل انسانی، هر دو به پیامدهای عاطفی دوزبانگی (اضطراب، کاهش اعتمادبه‌نفس و احساس شرمندگی) پرداختند، اما تحلیل انسانی ارتباط عمیق‌تری بین این تجربیات و عوامل زمینه‌ای (مانند سیاست‌های آموزشی و نگرش معلمان) برقرار کرده بود. تم «ناهمخوانی زبانی و تربیتی» در GPT-4 به‌صورت مستقل مطرح شد، درحالی‌که تحلیل انسانی آن را جزئی از چالش‌های نظام آموزشی در نظر گرفته بود. نکته قابل‌توجه، شناسایی تم «فقدان حمایت نهادی» توسط GPT-4 بود که اگرچه در تحلیل انسانی به‌صورت مستقل ظاهر نشده بود، اما نشان‌دهنده توانایی مدل در تشخیص الگوهای کلان در داده‌هاست. این تفاوت‌ها حاکی از آن است که GPT-4 در شناسایی اولیه مضامین دقیق عمل می‌کند، اما در ایجاد ارتباط بین مفاهیم و تفسیر عمیق‌تر زمینه‌های فرهنگی – اجتماعی محدودیت دارد.

اعتبارسنجی نتایج

به‌منظور ارزیابی اعتبار تحلیلی یافته‌های GPT-4، از روش آزمون تورینگ غیررسمی استفاده شد. آزمون تورینگ غیررسمی یعنی سنجش توانایی یک ماشین (برنامه رایانه‌ای) در تولید پاسخ‌هایی که به طور شهودی شبیه پاسخ انسان باشند، بدون آنکه لزوماً در قالب یک آزمون استاندارد تورینگ انجام شود (Jiang et al., 2024). در این آزمون، سه متخصص روش‌شناسی کیفی با سابقه پژوهش در حوزه آموزش دوزبانه، به‌صورت مستقل جدولی شامل ۱۰ تم را همراه با نقل‌قول‌های مرتبط، بدون اطلاع از منبع هر کدام، موردبررسی قرار دادند. نتایج نشان داد که دقت تشخیص متخصصان در تمایز بین تحلیل انسانی و ماشینی کمتر از ۲۰٪ بود که پایین‌تر از سطح تصادفی (۵۰٪) قرار دارد. این یافته نشان می‌دهد که از نظر متخصصان، خروجی GPT-4 از نظر انسجام مفهومی و ارتباط با داده‌ها، تفاوت معناداری با تحلیل پژوهشگران انسانی ندارد. بعلاوه برای سنجش پایایی درونی، آزمون ثبات اجرا شد. بدین منظور ترتیب متن مصاحبه‌ها به‌صورت تصادفی جابه‌جا گردید و داده‌ها مجدداً با همان پرامپت‌ها به GPT-4 ارائه شدند. در تحلیل مجدد، ۵ تم از مجموع ۶ تم اصلی دوباره شناسایی شد، که بیانگر ثبات نسبی در نتایج GPT-4 است.

بحث و نتیجه‌گیری

پژوهش حاضر با بهره‌گیری از رویکرد پدیدارشناسی توصیفی کلایزی و تحلیل از طریق زنجیره تفکر GPT-4، به شناسایی و تحلیل مشکلات آموزشی دانش‌آموزان دوزبانه پایه اول ابتدایی از دیدگاه معلمان پرداخت. دررابطه‌با پاسخ سؤال اول پژوهش مبنی بر آنکه مهم‌ترین مشکلاتی که دانش‌آموزان دوزبانه از منظر معلمان با آن مواجه‌اند، کدام‌اند؟ در تحلیل انسانی به روش کلایزی، چهار مقوله اصلی شامل مشکلات آموزشی، روان‌شناختی، اجتماعی و زبانی استخراج شد. در مقابل، تحلیل ماشینی مبتنی بر GPT-4، داده‌ها را در شش مضمون اصلی دسته‌بندی کرد که عبارت‌اند از: مشکلات زبانی پایه‌ای در درک و بیان زبان رسمی، تجربه‌های هیجانی و روانی منفی ناشی از فشار زبانی و فرهنگی، شکاف فرهنگی و احساس بیگانگی با محتوای آموزشی و محیط مدرسه، ناهمخوانی زبانی و تربیتی میان خانه و مدرسه، کناره‌گیری آموزشی و کاهش مشارکت فعال در کلاس، و فقدان حمایت ساختاری و بی‌توجهی نهادی به تفاوت‌های زبانی. این رویکرد ترکیبی، امکان مقایسه میان تحلیل انسانی و ماشینی و شناسایی ابعاد مکمل و هم‌پوشانی را فراهم ساخت. یافته‌های این پژوهش نشان می‌دهد که تحلیل انسانی و ماشینی هر کدام سهم منحصربه‌فردی در شناسایی

مشکلات دانش‌آموزان دوزبانه دارند. تحلیل انسانی با رویکرد کل‌نگر خود توانست ارتباط عمیق بین مشکلات زبانی، چالش‌های روان‌شناختی و نابرابری‌های آموزشی را در چارچوب نظریه‌های جامعه‌شناختی و روان‌شناسی تربیتی تبیین کند. این تحلیل نشان داد که زبان مادری صرفاً یک ابزار ارتباطی نیست، بلکه نقش حیاتی در شکل‌گیری هویت و فرایندهای شناختی دانش‌آموزان دارد. در مقابل، تحلیل ماشینی با شناسایی دقیق‌تر خطاهای زبانی و کشف تم‌های جدید مانند فقدان حمایت نهادی، ابعاد دیگری از مسئله را روشن ساخت که در تحلیل انسانی موردتوجه قرار نگرفته بود. در مجموع یافته‌های این پژوهش درباره تم‌های شناسایی شده با پژوهش‌های رحیم‌آبادی (۱۳۸۳)، کاظمی (۱۳۸۶)، عصاره (۱۳۸۷)، واحدی (۱۳۹۰)، فیاضی و همکاران (۱۳۹۳)، الله کرمی (۱۳۹۷)، الله کرمی و همکاران (۱۳۹۷) و هومین فر (Hoominfar, 2014) همسو است. یکی از مهم‌ترین یافته‌های این پژوهش، تم مشترک مشکلات زبانی بود که هم در تحلیل انسانی و هم در تحلیل ماشینی (GPT-4) شناسایی شد. در هر دو تحلیل، معلمان به چالش‌هایی مانند ضعف در درک زبان فارسی، ناتوانی در بیان مفاهیم درسی، و ترس از صحبت‌کردن اشاره کرده‌اند. این یافته با نظریه ناتوانی در زبان آکادمیک از جیم کامینز (Cummins, 2000) قابل‌تبیین است؛ دانش‌آموز ممکن است زبان محاوره را یاد گرفته باشد؛ اما در زبان علمی که در مدرسه استفاده می‌شود، دچار ضعف باشد. همچنین از منظر دیدگاه ویگوتسکی، زبان ابزاری کلیدی در یادگیری مفاهیم است و گسست میان زبان خانه و مدرسه موجب افت یادگیری می‌شود؛ بنابراین، نظام آموزشی باید ضمن احترام به زبان مادری، از آن به‌عنوان یک دارایی فرهنگی بهره گیرد و از راهکارهایی مانند آموزش دوزبانه، استفاده از رابط‌های زبانی در کلاس و طراحی محتوای قابل‌دسترس برای یادگیرندگان غیربومی استفاده کند. تم دوم، مشکلات روان‌شناختی و عاطفی بود که در هر دو تحلیل منعکس شده است. معلمان گزارش داده‌اند که دانش‌آموزان دوزبانه اغلب دچار اضطراب، خجالت، گوشه‌گیری و احساس طرد می‌شوند. تحلیل GPT-٤ نیز این مضامین را با اصطلاحاتی مانند اضطراب زبانی، ترس از تمسخر و احساس عدم تعلق منعکس کرد. از نظر نظری، این یافته با مدل اضطراب زبان‌آموزی (Horwitz et al., ١٩٨٦) و همچنین نظریه عزت‌نفس کوپر اسمیت هم‌راستا است. کودکانی که به دلیل ضعف زبانی مورد قضاوت یا تمسخر قرار می‌گیرند، از تعاملات اجتماعی و آموزشی دوری کرده و در فرایند یادگیری نیز انفعال نشان می‌دهند؛ بنابراین، کلاس‌های درس باید به فضاهایی ایمن، حامی و بدون قضاوت برای دانش‌آموزان دوزبانه تبدیل شوند و معلمان نیز

باید مهارت‌های تربیتی، عاطفی و فرهنگی را برای حمایت از این کودکان تقویت کنند. مشکلات فرهنگی و اجتماعی به‌عنوان سومین تم مشترک نیز در هر دو تحلیل برجسته بود. تحلیل‌های انسانی و ماشینی هر دو به تضاد میان ارزش‌های خانوادگی، باورهای فرهنگی، و زبان خانه با انتظارات رسمی مدرسه اشاره کرده‌اند. مفاهیمی مانند شکاف فرهنگی، نگرش منفی خانواده‌ها نسبت به مدرسه و بیگانگی کودک با فرهنگ مدرسه به‌خوبی منعکس‌کننده این مسئله‌اند. این یافته با نظریه سازگاری فرهنگی در آموزش (Gay & Hanley, 1999) قابل‌تحلیل است. هنگامی که زبان و فرهنگ دانش‌آموز در محیط مدرسه به رسمیت شناخته نمی‌شود، احساس تعلق کاهش یافته و مشارکت آموزشی افت می‌کند. بنابراین، ایجاد پیوند میان خانه و مدرسه از طریق گفت‌وگوی فرهنگی، مشارکت والدین، و طراحی آموزش حساس به زمینه فرهنگی، ضرورتی اساسی است. تم چهارم، مشکلات آموزشی و یادگیری نیز در هر دو تحلیل وجود داشت. معلمان به مواردی چون کندی یادگیری، نیاز به تکرار بیشتر، ضعف در حل مسئله، و ناتوانی در پیروی از ریتم کلاس اشاره کردند. تحلیل ماشینی نیز با عباراتی مانند بار شناختی بالا، پردازش آهسته اطلاعات، و وابستگی به تفسیر معلم این مسائل را شناسایی کرده است. این یافته با نظریه بار شناختی (Sweller, 2016) تبیین‌پذیر است؛ زیرا یادگیری از طریق زبان دوم ظرفیت حافظه فعال را کاهش داده و کارایی پردازش اطلاعات را پایین می‌آورد. در کنار این چهار تم مشترک، تحلیل ماشینی دو تم اضافی را نیز شناسایی کرد که در تحلیل انسانی مورد توجه قرار نگرفته بودند. نخستین آن، کناره‌گیری آموزشی و کاهش مشارکت فعال در کلاس بود. عدم پاسخ‌گویی به سؤالات، پرهیز از فعالیت داوطلبانه، و عدم مشارکت کارگروهی به‌عنوان نشانه‌های این کناره‌گیری تأکید کرده بود. این موضوع با نظریه‌های انگیزش در یادگیری، به‌ویژه نظریه خودتعیین‌گری دسی و رایان (Deci & Ryan, 1985) هم‌راستا است. زمانی که نیاز به شایستگی و تعلق در محیط یادگیری تأمین نشود، دانش‌آموز به‌جای درگیری شناختی، واکنش اجتنابی یا کناره‌گیرانه نشان می‌دهد. دومین تم اختصاصی تحلیل ماشینی، نیاز به سیاست‌گذاری نهادی و زبانی مؤثر بود. تحلیل ماشین به کاستی‌های ساختاری مانند نبود برنامه‌های ملی حمایت زبانی، کمبود معلمان آموزش‌دیده در حوزه دوزبانه، و ضعف نظام ارزشیابی متناسب با نیازهای زبانی دانش‌آموزان اشاره کرده بود. می‌توان گفت اگر ساختارهای کلان آموزشی برای شناسایی و حمایت از تنوع زبانی طرحی نداشته باشد، نابرابری آموزشی بازتولید خواهد شد؛ بنابراین، نظام آموزشی نیازمند بازنگری در سیاست‌گذاری زبانی، طراحی برنامه‌های درسی با

درنظرگرفتن نقش زبان مادری، و تربیت معلمان توانمند در زمینه آموزش دوزبانه است تا بتواند آموزش را از انحصار زبانی خارج کرده و به سمت عدالت و برابری واقعی سوق دهد.

در پاسخ به سؤال دوم پژوهش، که به مقایسه تحلیل مشکلات دانش‌آموزان دوزبانه با استفاده از روش پدیدارشناسی کلایزی و تحلیل همان داده‌ها به‌وسیلهٔ مدل زبانی GPT-4 مبتنی بر راهبرد زنجیره تفکر می‌پردازد، می‌توان بیان کرد که تحلیل مبتنی بر هوش مصنوعی، به‌ویژه به‌دلیل توانمندی در شناسایی خطاهای ساختاری و کشف تم‌های نو، ابعاد متفاوتی از مسئله را برجسته ساخت. مقایسه تم‌های استخراج‌شده از دو روش نیز یافته‌های قابل‌تأملی دربرداشت. از جمله، سوگیری‌های ناخودآگاه تحلیلگران انسانی که ممکن است تحت تأثیر گفتمان رسمی و سیاسی، برخی مسائل ساختاری را نادیده گرفته یا به شکل ناخودآگاه آن‌ها را عادی‌سازی کرده باشند. این در حالی است که مدل‌های زبانی، به‌دلیل توانایی منحصربه‌فرد در شناسایی الگوهای فرامتنی و روابط پنهان در داده‌ها، قادر به استخراج لایه‌های عمیق‌تری از معنا بوده‌اند. این تفاوت‌ها نشان می‌دهد که بهره‌گیری هم‌زمان از تحلیل انسانی و ماشینی می‌تواند به غنای فهم پدیدهٔ مورد مطالعه بیفزاید.

با بررسی دقیق‌تر مصاحبه‌ها، می‌توان دریافت که معلمان به‌صورت مستقیم به سیاست‌های کلان اشاره نکرده‌اند، اما نشانه‌های غیرمستقیمی از کمبود منابع آموزشی، محدودیت‌های سیاستی و نبود برنامه‌ریزی درسی چندفرهنگی در اظهارات آنها قابل‌تشخیص است. این غفلت تحلیلی می‌تواند پیامدهای مهمی داشته باشد، از جمله کاهش سطح تحلیل به ابعاد فردی و تضعیف رویکرد انتقادی در پژوهش. از منظر روش‌شناسی کیفی، مقایسه بین تحلیل انسانی و ماشینی نشان‌دهنده تفاوت‌های بنیادین در سطح تحلیل، عمق تفسیر و چارچوب‌های مفهومی است. در حوزه مسائل زبانی، تحلیل انسانی با رویکردی کل‌نگر، چالش‌های زبانی را در چارچوب گسترده‌تری از مشکلات آموزشی موردبررسی قرار داد که شامل ارتباط پویای زبان، فرایند یادگیری و محتوای درسی بود. این رویکرد که ریشه در نظریه‌های اجتماعی – فرهنگی آموزش مانند نظریه ویگوتسکی دارد، توانست پیوندهای معناداری بین مشکلات زبانی و بستر آموزشی ایجاد کند. در مقابل، تحلیل ماشینی GPT-4 بیشتر بر جنبه‌های صوری زبان مانند خطاهای ساختاری و دستوری تمرکز داشت و نتوانست ارتباط نظام‌مند این چالش‌ها با فرایندهای یادگیری را به طور کامل تبیین کند. در حوزه پیامدهای روانی، تحلیل انسانی با تلفیق مفاهیم اضطراب، اعتمادبه‌نفس و انزوای اجتماعی در چارچوب چالش‌های روان‌شناختی، توانست وحدت تجربه زیسته معلمان و دانش‌آموزان را بازنمایی

کند. این رویکرد که مبتنی بر پدیدارشناسی توصیفی کلایزی است، پیچیدگی‌های تجربه دوزبانگی را در بافت فرهنگی اورامانات به‌خوبی بازتاب داد. در مقابل، GPT-4 این مفاهیم را به تم‌های جداگانه «تجربه‌های هیجانی منفی» و «کناره‌گیری آموزشی» تقسیم کرد و نتوانست ارتباط درونی این پدیده‌ها را که در داده‌ها مشهود بود، تشخیص دهد. این محدودیت ناشی از ناتوانی مدل‌های زبانی در درک ارتباطات معنایی عمیق و تفسیر بافتارهای فرهنگی پیچیده است. در تحلیل تفاوت‌های تعداد و ماهیت تم‌ها، سه عامل کلیدی نقش دارد: اولاً، تحلیل انسانی در پدیدارشناسی به دنبال کشف وحدت معنایی است، تم‌های مرتبط را در چارچوب‌های کلی‌تر ادغام می‌کند؛ درحالی‌که GPT-4 با الگوریتم‌های آماری، به تفکیک دقیق مفاهیم بر اساس شاخص‌های کمّی مانند تکرار واژگان می‌پردازد. ثانیاً، تحلیل انسانی از چارچوب‌های نظری مانند نظریه هویت اجتماعی یا نظریه‌های انتقادی آموزش برای تفسیر داده‌ها بهره می‌برد، حال‌آنکه GPT-4 فاقد چنین چارچوب‌های مفهومی است. ثالثاً، حساسیت به بافت فرهنگی در تحلیل انسانی منجر به تفسیرهای غنی‌تر و چندبعدی می‌شود، درحالی‌که تحلیل ماشینی عمدتاً به سطح ظاهری داده‌ها محدود می‌ماند؛ بنابراین، این مقایسه نشان می‌دهد که هرچند مدل‌های زبانی مانند GPT-4 در شناسایی اولیه مضامین و دسته‌بندی داده‌های حجیم مفید هستند، اما برای تحلیل‌های عمیق کیفی که نیازمند درک ظرایف فرهنگی، ارتباطات معنایی پیچیده و تفسیر نظری هستند، باید به‌عنوان مکمل تحلیل انسانی به کار گرفته شوند. ترکیب این دو رویکرد، یعنی استفاده از GPT-4 برای شناسایی اولیه تم‌ها و تحلیل انسانی برای تفسیر عمیق و یکپارچه‌سازی نظری، می‌تواند راهبرد مناسبی برای پژوهش‌های کیفی آینده باشد. رویکردی ترکیبی در تحلیل داده‌ها، درک عمیق‌تری از موضوع فراهم می‌کند و می‌تواند مکمل تحلیل‌های سنتی باشد. در همین راستا، یافته‌های این پژوهش با نتایج مطالعات کرسول و کرسول (Creswell & Creswell, ۲۰۱۷) و اویانگ و همکاران (Ouyang et al, 2022) همسو است؛ مطالعاتی که نشان می‌دهد ترکیب این دو رویکرد می‌تواند به درک جامع‌تری از مسائل پیچیده آموزشی منجر شود. بااین‌حال، تحلیل ماشینی چالش‌های مهمی نیز به همراه دارد؛ از جمله سوگیری‌های احتمالی (Zhang et al, 2018)، محدودیت در درک زمینه فرهنگی (Denzin & Lincoln, 2011) و نیاز به شفافیت روش‌شناختی (Grimmer et al, 2022). براین‌اساس، همان‌گونه که بویدا و آناما (Boveda & Annamma, 2023) تأکید کرده‌اند، به‌کارگیری هوش مصنوعی در پژوهش‌های کیفی باید به‌عنوان مکمل تحلیل انسانی و با حفظ «حضور

انسان در چرخه تحلیل» انجام شود تا هم از مزایای فناوری بهره برده و هم از محدودیت‌های آن در امان ماند. از منظر روش‌شناسی، مقایسه دو روش تحلیل، تفاوت‌های بنیادینی را نشان می‌دهد. نخست اینکه، بی‌طرفی کامل در تحلیل ماشینی قابل تضمین نیست، زیرا نتایج آن به‌شدت به نوع پرامپت‌دهی، ترتیب و نحوه‌ی ارائه داده‌ها وابسته است. دوم، چالش تفسیر تم‌ها در پژوهش کیفی ماهیتی ذهنی دارد و بدون چارچوب تحلیلی مشترک، تمایز یا هم‌پوشانی مفهومی بین دسته‌بندی‌های دو روش ممکن است ابهام‌برانگیز باشد. همچنین، نبود معیار مشخص برای اعتباربخشی نتایج، موجب ابهام در ارزیابی تفاوت‌های احتمالی میان یافته‌های دو روش می‌شود و این نیازمند داوری ثالث یا ابزار ارزیابی دقیق‌تری است. از سوی دیگر، باید به این نکته توجه کرد که GPT-4 ممکن است در پردازش واژگان محلی یا سبک گفتاری خاص منطقه اورامانات، عملکرد ضعیف‌تری داشته باشد. هر چند در پژوهش حاضر شباهت قابل‌توجه به تحلیل انسانی و ثبات تحلیلی بالا اهمیت استفاده از ماشین در تحلیل ها را نشان می‌دهد؛ بااین‌حال، تفاوت‌های موجود در تفسیر بافت‌های فرهنگی، لزوم تلفیق این ابزار با قضاوت تخصصی پژوهشگران را خاطرنشان می‌سازد. این یافته با مطالعات مشابه در حوزه پردازش زبان طبیعی (Jiang et al., 2024; Hamilton et al, 2023) همسو است که بر قابلیت‌های مدل‌های زبانی بزرگ در تحلیل‌های ساختاریافته تأکید دارند. از لحاظ زمان پردازش نیز بین دو روش تحلیل تفاوت چشمگیری وجود دارد GPT-4 کمتر از یک دقیقه صرف کرد تا کل فرایند از پردازش متن مصاحبه تا تفسیر نهایی را انجام دهد، درحالی‌که تحلیل انسانی داده‌ها حدود ۱۷ ساعت زمان برد. نکته مهم این است که GPT-٤ توانایی مقیاس‌پذیری بالایی دارد و می‌تواند دو تا سه برابر این حجم داده را در همان زمان پردازش کند، ولی پردازش انسانی برای حجم مشابه داده به زمان بسیار بیشتری نیاز دارد.

نتایج این پژوهش دو پیامد مهم برای پژوهش‌های کیفی دارد: نخست آنکه GPT-4 می‌تواند به عنوان ابزاری کارآمد در فرایندهای اولیه تحلیل کیفی، مانند کدگذاری داده‌ها یا شناسایی تم‌های اصلی، مورداستفاده قرار گیرد. این ویژگی به‌ویژه در مطالعات اکتشافی با حجم داده‌های زیاد می‌تواند مفید باشد. دوم آنکه علی‌رغم این شباهت ظاهری، باید توجه داشت که ممکن است تفاوت‌های ظریفی در عمق تحلیل، توجه به جزئیات بافت‌محور و تفسیرهای نظری بین تحلیل انسانی و ماشینی وجود داشته باشد؛ بنابراین، استفاده از GPT-4 در پژوهش‌های کیفی حساس به بافت فرهنگی، مانند مطالعه حاضر، باید بااحتیاط و همراه با نظارت و قضاوت پژوهشگران انسانی صورت پذیرد. بعبارتی مدل‌های زبانی بزرگ فاقد

آگاهی، تجربه زیسته یا قصدیت معنادار هستند و به‌جای تجربه، الگوها و احتمالات آماری در متون را بازتولید می‌کنند. بنابراین هر خروجیِ ماشین صرفاً یک سازوکار آماری–زبانی برای کشف همایندی‌ها و پیشنهاد تم‌های اولیه است و نمی‌تواند جایگزین تفسیر پدیدارشناسانه انسانی باشد. به همین دلیل، باید خروجی‌های ماشینی فقط به‌عنوان پیشنهادات اولیه وارد فرایند شود و تفسیر نهایی و استخراج ساختار بنیادین تجربه توسط پژوهشگران انسانی انجام گیرد. در پایان، این پژوهش پیشنهاد می‌کند که مطالعات آینده با ترکیب روش‌های کیفی و کمّی، استفاده از نمونه‌های گسترده‌تر، و بهره‌گیری از قابلیت‌های تحلیلی پیشرفته‌تر هوش مصنوعی، به توسعه مدل‌های آموزشی مناسب‌تری برای دانش‌آموزان دوزبانه کمک کنند. همچنین، انجام پژوهش‌های طولی برای بررسی تأثیر برنامه‌های آموزشی دوزبانه بر پیشرفت تحصیلی و رشد روانی - اجتماعی این دانش‌آموزان ضروری به نظر می‌رسد.

ملاحظات اخلاقی

نویسندگان اصول اخلاقی را در انجام و انتشار این پژوهش علمی رعایت نموده‌اند و این موضوع مورد تأیید همه آنهاست.

مشارکت نویسندگان

همهٔ نویسندگان به‌یکسان در مفهوم‌سازی مقاله و نگارش نسخهٔ اولیه و نسخه‌های بعدی آن مشارکت داشته‌اند.

تعارض منافع

نویسندگان هیچ‌گونه تعارض منافعی ندارند.

حامی مالی

مقاله حاضر بدون حمایت مالی وبه صورت مستقل توسط نویسندگان انجام شده است.

سپاسگزاری

نویسندگان مراتب سپاس خود را از همهٔ مشارکت‌کنندگان این پژوهش، به‌ویژه معلمان دورهٔ ابتدایی که با اشتراک‌گذاری ارزشمند دیدگاه‌های خود، زمینهٔ غنای این مطالعه را فراهم آوردند، ابراز می‌دارند.

منابع

اسپادا، نینا و لایت باون، پاتسی. (۱۳۹۵). *زبان‌ها چگونه آموخته می‌شوند*. (ترجمه رضا مراد صحرائی). تهران: انتشارات دانشگاه علامه طباطبائی.

الله کرمی، آزاد. (۱۳۹۷). ارائه الگوی آموزش مبتنی بر زبان مادری با رویکرد شناسایی مشکلات دانش‌آموزان دوزبانه و بررسی تأثیر آن بر مفهوم‌آموزی. رساله دکتری، دانشگاه علامه طباطبائی تهران.

الله کرمی، آزاد، امیری، حجت، و طوفانی نژاد، احسان. (۱۳۹۶). چندرسانه ای تعاملی، سواد قرآنی و انگیزش تحصیلی. *فناوری آموزش و یادگیری*، ۳(۱۲)، ۲۷-۴۶. https://doi.org/10.22054/jti.2020.1024.1029

الله کرمی، آزاد؛ علی آبادی، خدیجه؛ صحرایی، رضا مراد و دلاور، علی. (۱۳۹۷). آموزش به زبان مادری در نظام آموزش رسمی کشور: فرصت یا تهدید؟. *فصلنامه نوآوری‌های آموزشی*. ۶۵(۱۷). ۷۴-۴۳. https://noavaryedu.oerp.ir/article_79139.html?lang=en

رحیم آبادی، رحمت اله. (۱۳۸۳). *مقایسه عملکرد تحصیلی دانش‌آموزان یک‌زبانه و دو زبانه در دوره ابتدایی*. پایان نامه کارشناسی ارشد، دانشگاه آزاد اسلامی‌واحد رودهن.

سبحانی سنجبد، حاتم. (۱۳۹۵). *نقش زبان مادری در آموزش*. تبریز: نشر اختر.

صدری‌دمیرچی، اسماعیل، عباسی، ناصر، قهرمانلو، مریم، قربانی نژاد، علی و نوروزی همایون، محمدرضا . (۱۴۰۴). ویژگی‌های روانسنجی مقیاس نگرش کلی به هوش مصنوعی. *فصلنامه اندازه گیری تربیتی. ۱۶(۶۱)، ۱۵۹-۱۸۱*. doi: 10.22054/jem.2025.82349.3576

عصاره ، فریده. (۱۳۸۷). بررسی مشکلات زبان‌آموزی کودکان پایه ی اول ابتدایی در مناطق دو زبانه. *مجله مطالعه برنامه درسی*، شماره ۹، ۴۲-۶۰.

فیاضی، علی، صحراگرد، رحمان؛ بلقیس، روشن و زندی، بهمن. (۱۳۹۳). مقایسه دانش زبان انگلیسی دانش‌آموزان دوزبانه استان خوزستان و دانش‌آموزان یک‌زبانه استان فارس در مقطع چهارم متوسطه: رویکرد شناختی. *مجله زبان شناسی و گویش های خراسان*. شماره پیاپی ۱۰. ص ۹۶-۷۰. https://doi.org/10.22067/lj.v6i10.36574

قدم پور، عزت اله، جعفری, مسعود، امامی، شیرین و بهرامی، عرفان. (۱۴۰۴). ویژگی‌های روانسنجی مقیاس سواد هوش مصنوعی در دانشجویان ایرانی. *فصلنامه اندازه گیری تربیتی، ۱۶(۶۲)، ۱۲۹-۱۵۷*. doi: 10.22054/jem.2025.84551.3600

کاظمی، محمد رضا. (۱۳۸۶). *بررسی تأثیر دوره یک ماهه آمادگی دو زبانه بر زبان آموزی و درس فارسی دانش‌آموزان پایه اول ابتدایی مناطق دو زبانه شهرستان های اهواز و رامهرمز*. پایان نامه کارشناسی ارشد، دانشگاه علامه طباطبایی تهران.

نرسیسیانس، امیلیا (۱۳۸۲). *مباحثی در دوزبانگی از دیدگاه علوم اجتماعی*. تهران: سازمان میراث فرهنگی کشور، نشر بقعه.

واحدی، شهرام (۱۳۹۰). *رویکردها و استراتژی های آموزشی دو زبانه و پیامدهای شناختی عاطفی و اجتماعی آن*. در اسداله، خدیوی؛ رضا، کلانتری و رجب ربوی(گردآورندگان)، (۱۳۹۰). مجموعه مقالات همایش دوزبانگی و آموزش: چالش‌ها، چشم اندازها و راهکارها(ص۱۲۲-۱۵۵). تهران، پژوهشگاه مطالعات آموزش‌وپرورش.

References

Alabi, A.T. (2002). Using Yoruba as a Medium of Instruction in the Teaching of Primary Science": An Evaluation of the Implementation of the National Policy on education. *Journal of Sokoto Educational Review*, 5: pp 44 - 57.

Albert, R. T., Albert, R. E., & Radsma, J. (2002). Relationships among bilingualism, critical thinking ability, and critical thinking disposition. *Journal of Professional Nursing*, 18(4), 220-229.

Allah Karami,A , Aliābādi,K , Sahrāyi,R and delavar,A . (2018). Education in mother tongue in the formal education of Iran: Opportunity or threat?. *Journal of Educational Innovations*, *17*(1), 43-74. https://noavaryedu.oerp.ir/article_79139.html?lang=en. [In Persian].

Allahkarami, A. (2018). Providing a Mother Tongue-Based Education Model with an Approach to Identifying Bilingual Students' Problems and Examining Its Effect on Concept Acquisition. Doctoral dissertation, Allameh Tabataba'i University, Tehran. [In Persian].

Asareh, F. (2008). An investigation into language learning problems of first-grade elementary students in bilingual areas. Journal of Curriculum Studies, 9, 42-٦٠. [In Persian].

Baker, C. (2008). *Foundations of bilingual education and bilingualism.* Clevedon: Multilingual Matters.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021).On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922.

Bengio, Y., Ducharme, R., & Vincent, P. (2000). A neural probabilistic language model. *Advances in Neural Information Processing Systems*, *33*, 1877–1901.

Benjamin, S., Williams, J., & Maher, M.A. (2017). Focusing the lens to share the story: Using photographs and interviews to explore doctoral

students' sense of well-being. *International Journal of Doctoral Studies*, *12*(1), 197–217.
Bhardwaj, R., Majumder, N., & Poria, S. (2021). Investigating Gender Bias in BERT. *Cognitive Computation*, *13*(4), 1008–1018. https://doi.org/10.1007/s12559-021-09881-2
Bialystok, E. (2007). Cognitive effects of bilingualism: How linguistic experience leads to cognitive change. *International Journal of Bilingual Education and Bilingualism*, 10(3), 210-223. *bias in BERT*. arXiv. https://doi.org/10.48550/arXiv.2009.05021
biegbu, I. (2016). Mother tongue education: a panacea to effective teaching and learning of English in Nigeria primary schools. *AFRREV LALIGENS: An International Journal of Language, Literature and Gender Studies*, *5*(2), 50–59. https://doi.org/10.4314/LALIGENS.V5I2.5
Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. *Proceedings of the National Academy of Sciences (PNAS)*, *120*(6), Article e2218523120. https://doi.org/10.1073/pnas.2218523120.
Birhane, A., van Dijk, J., & Andreassen, P. B. (2022). Power to the people? Opportunities and challenges for participatory approaches to AI development. *Patterns*, 3(10), 100606. https://doi.org/10.1016/j.patter.2022.100606
Boveda, M., & Annamma, S. A. (2023). Beyond making a statement:An intersectional framing of the power and possibilities of positioning.*Educational Researcher*, *52*(5), 306–314. https://doi.org/10.3102/0013189X231167149
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D., Wu, J., Winter, C., . . . Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*, 1877–1901. https://proceedings
Chowdhary, K. (2020). *Fundamentals of artificial intelligence*. Springer. https://doi.org/10.1007/978-81-322-3972-7_19
Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative,quantitative, and mixed methods approaches*. SAGE Publications.
Cummins, J. (2000). *Bilingual children's mother tongue: Why is it important for education?* Sprogforum, vol/issue: 7 (19), pp. 15-20. Retrieved March 26, 2016 from http://www.iteachilearn.com/cummins/mother.htm.
Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum Press.
Denzin, N. K., & Lincoln, Y. S. (2011). *The SAGE handbook of qualitative research* (5th ed.). SAGE Publications.

Devlin, J., Chang, M.W., Lee, K., et al. (2018) Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. arXiv: 1810.04805. https://doi.org/10.48550/arXiv.1810.04805

Diaz, R. M. (1985). Bilingual cognitive development: Addressing three gaps in current research. *Child development*, 1376-1388.

Diba, F., & Priyadarshini, A. (2023). Mother Tongue Education as a Medium of Instruction - A Literature Review. *International Journal For Multidisciplinary Research*. https://doi.org/10.36948/ijfmr.2023.v05i06.9679.

Elkins, K., & Chun, J. (2020). Can GPT-3 pass a writer's Turing test? *Journal of Cultural Analytics, 5*(2). https://doi.org/10.22148/001c.17212

Fayyazi, A. , Sahragard, R. , Rovshan, B. and Zandy, B. (2014). A Comparative Study of Bilingual and Monolingual Iranian EFL students' Knowledge of English Grammar and Vocabulary: A Cognitive Approach. *Journal of Linguistics and Khorasan Dialects*, *6*(10), 69-96. https://doi.org/10.22067/lj.v6i10.36574. [In Persian].

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, *30*(4), 681–694. https:// doi.org/10.1007/s11023-020-09548-1

Gay, G., & Hanley, M. S. (1999). Multicultural Empowerment in Middle School Social Studies through Drama Pedagogy. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, *72*(6), 364–370. https://doi.org/10.1080/00098659909599426

Ghadampour, E. , jafari, M. , Emami, S. and bahrami, E. (2025). The psychometric properties of the Artificial Intelligence Literacy Scale (AILS-CCS) in Iranian students.. *Quarterly of Educational Measurement*, *16*(62), 129-157. doi: ١٠,٢٢٠٥٤/jem.٢٠٢٥,٨٤٥٥١,٣٦٠٠. [In Persian].

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. PrincetonUniversity Press.

Gunning, C. S. (1981). *Relationships among field independence, critical thinking ability, and clinical problem-solving ability of baccalaureate nursing students* (Doctoral dissertation, University of Texas at Austin).

Hamilton, L., Elliott, D., Quick, A., Smith, S., & Choplin, V. (2023). Exploring the use of AI in qualitative analysis: A comparative study MONTH XXXX 11of guaranteed income data. *International Journal of Qualitative Methods*, *22*, 1–13. https://doi.org/10.1177/16094069231201504

Hoominfar, E. (2014). Challenges of Monolingual Education. (Doctoral dissertation, Bowling Green State University).

Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *Tesol Quarterly*, *20*(3), 559-562.

Isoaho, K., Gritsenko, D., & Mäkelä, E. (2021). Topic modeling and text analysis for qualitative policy research. *Policy Studies Journal*, *49*(1), 300–324. https://doi.org/10.1111/psj.12343

Jacinto, A.L. (2013). The local dialect as subject and medium of instruction. *The Modern Teacher*. Vol. 62, No. 3, pp. 85.

Jiang, Y., Ko-Wong, L., & Valdovinos Gutierrez, I. (2025). The Feasibility and Comparability of Using Artificial Intelligence for Qualitative Data Analysis in Equity-Focused Research. Educational Researcher, 54(3), 153-163. https://doi.org/10.3102/0013189X251314821

Kazemi, M. R. (2007). Investigating the Effect of a One-Month Bilingual Preparatory Course on Language Learning and Persian Lessons of First-Grade Elementary Students in Bilingual Areas of Ahvaz and Ramhormoz Counties. Master's thesis, Allameh Tabataba'i University, Tehran. [In Persian].

Konaka, K. (1997). *The relationship between degree of bilingualism and gender to divergent thinking ability among native Japanese-speaking children in New York area*. Doctoral dissertation, New York University, 1997. Dissertation Abstracts International, 58, AAM9718715.

Malmberg, B. (1972). Descriptive linguistics. *Current Trends in Linguistics,* 9(1), 223-253.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. arXiv. https://doi.org/10.48550/arXiv.1301.3781

Nersisians, E. (2003). Topics in Bilingualism from a Social Science Perspective. Tehran: Cultural Heritage Organization of Iran, Bogheh Publications. [In Persian].

Nieto, S. (2000). *Affirming diversity: The sociopolitical context of multicultural education.* New York: Longman.

Ouedraogo, R. M. (2000). *Language planning and language policies in some selected West African countries*. International Institute for Capacity Building in Africa.

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P. F., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, *35*, 27730–27744.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, (2019). Language models are unsupervised multitask learners. *OpenAI blog*, *1*(8), 9.

Rahimabadi, R. (2004). Comparison of Academic Performance of Monolingual and Bilingual Students in Elementary School. Master's thesis, Islamic Azad University, Roudhen Branch. [In Persian].

Sadri Damirchi, E. , Abbasi, N. , Ghahremanloo, M. , Ghorbaninejad, A. and Noroozi Homayoon, M. (2025). Psychometric Properties of the General Attitude Toward Artificial Intelligence Scale. *Quarterly of Educational Measurement*, ۱۶(۶۱), ۱۵۹-۱۸۱. doi: ۱۰,۲۲۰۵٤/jem.۲۰۲۵,۸۲۳٤۹,۳۵۷٦. [In Persian].

Skutnabb-Kangas, T. (1981). *Bilingualism or not: The education of minorities* (Vol. 7). Multilingual matters.

Sobhani Sanjbad, H. (2016). The Role of Mother Tongue in Education. Tabriz: Akhtar Publications. [In Persian].

Spada, N., & Lightbown, P. (2016). How Languages Are Learned. (Translated by Reza Morad Sahraei). Tehran: Allameh Tabataba'i University Press. [In Persian].

Sweller, J. (2016). Working memory, long-term memory, and instructional design. *Journal of Applied Research in Memory and Cognition*, *5*(4), 360-367.

Titone, R. (2013). Psychological aspects of multilingual education. *International Review of Education*, 24(3). 283-293.

Vahedi, S. (2011). Bilingual educational approaches and strategies and their cognitive, affective, and social consequences. In A. Khadivi, R. Kalantari, & R. Rabavi (Eds.), Proceedings of the Conference on Bilingualism and Education: Challenges, Prospects, and Solutions (pp. 122-155). Tehran: Institute for Educational Studies. [In Persian].

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. arXiv. https://doi.org/10.48550/arXiv.1706.03762

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2022). *Chain-of-thought prompting elicits reasoning in large language models*. arXiv. https://doi.org/10.48550/arXiv.2201.11903

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P. S., ... & Gabriel, I. (2021). Ethical and social risks of harm from Language Models. *arXiv preprint*, arXiv:2112.04359. https://doi.org/10.48550/arXiv.2112.04359

Ye, J., Chen, X., Xu, N., Zu, C., Shao, Z., Liu, S., Cui, Y., Zhou, Z., Gong, C., Shen, Y., Zhou, J., Chen, S., Gui, T., Zhang, Q., & Huang, X. (2023). *A comprehensive capability analysis of gpt-3 and gpt-3.5 series models*. arXiv. https://doi.org/10.48550/arXiv.2303.10420

Zhang, B. H., Lemoine, B., & Mitchell, M. (2018). Mitigating unwanted biases with adversarial learning. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 335–340). Association for Computing Machinery. https://doi.org/10.1145/3278721.3278779