

Article

Enhancing IEP Design in Inclusive Primary Settings Through ChatGPT: A Mixed-Methods Study with Special Educators

Stergiani Giaouri ^{1,*}  and Maria Charisi ² ¹ Department of Early Childhood Education, School of Social Sciences and Humanities, University of Western Macedonia, 53100 Florina, Greece² Department of Primary Education, School of Social Sciences and Humanities, University of Western Macedonia, 53100 Florina, Greece; edumal01134@uowm.gr

* Correspondence: sgiaouri@uowm.gr

Abstract

The integration of Artificial Intelligence (AI) in education has raised important questions about its role in supporting inclusive practices, particularly in special education. This qualitative-dominant study with quantitative support examines how special education teachers in inclusive primary classrooms in Greece use ChatGPT to design Individualized Education Programs (IEPs) for students with learning disabilities. Six teachers participated, with some employing ChatGPT and others relying on traditional methods. The quality of IEP goals was described using the Revised IEP/IFSP Goals and Objectives Rating Instrument (R-GORI), while in-depth teacher perspectives were explored through thematic analysis. Findings suggest that ChatGPT contributed to clearer goal-setting, generation of diverse instructional resources, and more structured lesson planning. However, teachers emphasized the need for critical oversight, adaptation to real-world classroom conditions, and safeguarding the relational and emotional aspects of teaching. Participants expressed cautious optimism, viewing ChatGPT as a valuable support tool when integrated thoughtfully and ethically. These context-specific, exploratory results offer preliminary guidance for educators, policymakers, and researchers seeking to integrate AI tools into special education. They highlight the importance of targeted professional development, ethical safeguards, and further large-scale research to evaluate the broader applicability of AI-assisted IEP planning.

Keywords: ChatGPT; individualized education programs; special and inclusive education; learning disabilities; primary education; teacher practices



Academic Editors: Aglaia Stampoltzis and Eleni Peristeri

Received: 25 July 2025

Revised: 14 August 2025

Accepted: 18 August 2025

Published: 19 August 2025

Citation: Giaouri, S., & Charisi, M. (2025). Enhancing IEP Design in Inclusive Primary Settings Through ChatGPT: A Mixed-Methods Study with Special Educators. *Education Sciences*, 15(8), 1065. <https://doi.org/10.3390/educsci15081065>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The integration of Artificial Intelligence (AI) into education has attracted growing interest, particularly with the advent of large language models such as ChatGPT. As a generative AI tool, ChatGPT can produce coherent text, summaries, translations, creative content, and even images in response to user input (Eke, 2023; Rospigliosi, 2023). Its versatility has made it increasingly attractive for use in educational contexts, where it supports a variety of instructional and administrative tasks (Lund & Wang, 2023).

Over the past two years, interest in the educational potential of ChatGPT has grown rapidly. In primary education, studies by Baytak (2024), Karaman and Göksu (2024), Lee and Zhai (2024), Rashed Ibraheam Almohesh (2024), Yilmaz Can and Durmuş (2024), and Zemljak (2023) have examined its use across various teaching and learning contexts.

Research in secondary education is less common, with work by [Khuibut et al. \(2023\)](#), [Küchemann et al. \(2023\)](#), and [Li et al. \(2024\)](#) offering initial insights. Within special education, investigations remain scarce, though contributions from [Rakap \(2023\)](#), [Rakap and Balikci \(2024\)](#), [Rizos et al. \(2024\)](#), and [Waterfield et al. \(2025\)](#) highlight emerging possibilities. Geographically, the majority of studies have been conducted in Turkey and Europe, while interest is beginning to grow in Asia, the Middle East, and the United States.

From a methodological perspective, recent research has adopted a range of approaches. Several studies—such as those by [Karaman and Göksu \(2024\)](#), [Khuibut et al. \(2023\)](#), [Küchemann et al. \(2023\)](#), [Rakap \(2023\)](#), [Rakap and Balikci \(2024\)](#), [Rashed Ibraheam Almohesh \(2024\)](#), and [Yilmaz Can and Durmuş \(2024\)](#)—have used experimental designs with control and intervention groups. Others, including [Baytak \(2024\)](#), [Lee and Zhai \(2024\)](#), [Li et al. \(2024\)](#), [Waterfield et al. \(2025\)](#), and [Zemljak \(2023\)](#), have relied on single-group designs in which all participants used ChatGPT. In terms of subject focus, most applications have been in STEM fields, such as mathematics, physics, and biology, as seen in the work of [Karaman and Göksu \(2024\)](#), [Lee and Zhai \(2024\)](#), [Rizos et al. \(2024\)](#), [Yilmaz Can and Durmuş \(2024\)](#), and [Zemljak \(2023\)](#). Far fewer have explored the humanities, with [Baytak \(2024\)](#) and [Li et al. \(2024\)](#) providing some examples.

In the context of special education, [Rakap \(2023\)](#) examined the impact of ChatGPT on the quality, content, and efficiency of IEP goal development by novice special educators. Using the Revised IEP/IFSP Goals and Objective Rating Instrument (R-GORI) ([Notari-Syverson & Shuster, 1995](#)), the study found measurable benefits in goal formulation. A follow-up study by [Rakap and Balikci \(2024\)](#) further emphasized the added value of ChatGPT, particularly when combined with targeted training in SMART goal-setting ([Jung, 2007](#)). Similarly, [Rizos et al. \(2024\)](#) implemented a ChatGPT-supported intervention for two students with dyslexia and autism spectrum disorder (ASD) in Greece, demonstrating the potential of personalized AI-generated materials in mathematics instruction. In a broader U.S.-based study, [Waterfield et al. \(2025\)](#) involved 56 special educators across 22 states, comparing independently written IEP goals with those co-developed using ChatGPT. These studies collectively highlight both the promise and the practical challenges of integrating generative AI into individualized educational planning.

Despite encouraging findings, the current body of research on the use of ChatGPT in education remains inconclusive. Variability in educators' expertise, subject areas, student profiles, and the specific versions of ChatGPT employed may contribute to the inconsistency of results. Furthermore, limited research has examined how AI tools interact with factors such as teachers' prior training, time investment, and professional perceptions—particularly within inclusive classrooms that serve students with learning disabilities (LDs). [Adel et al. \(2024\)](#) investigate the pedagogical potential and ethical challenges of ChatGPT, highlighting its computational benefits for teaching and learning, while also raising important concerns regarding academic integrity, data protection, and responsible AI implementation. Similarly, [Pagliara et al. \(2024\)](#), through a comprehensive scoping review, examine the integration of AI in inclusive education, emphasizing its capacity to support personalized learning. Their findings stress the importance of clear ethical guidelines, targeted teacher training, and equitable access to digital tools. Complementing these perspectives, [Van den Berg and du Plessis \(2023\)](#) explore how ChatGPT and generative AI can contribute to lesson planning, critical thinking, and openness in teacher education. They underscore the value of reflective practice and argue that meaningful AI integration requires strong pedagogical grounding and ethical awareness.

The present study addresses a critical gap in the literature by investigating the use of ChatGPT 4o-mini in the design of Individualized Education Programs (IEPs) for students with learning disabilities (LDs) by special education teachers in inclusive primary

classrooms in Greece. The primary aim is to explore teacher perceptions, experiences, and practical considerations in integrating ChatGPT into the IEP development process. Adopting a qualitative-dominant design with quantitative support, the research prioritizes in-depth thematic analysis of teacher narratives while using descriptive R-GORI scores to illustrate patterns in IEP goal quality. The quantitative component is limited by the very small sample size and is therefore not intended for statistical generalization, serving instead to complement the qualitative findings. The study also considers how prior training in IEP development and the amount of planning time may relate to observed variations in goal quality. These context-specific insights aim to inform best practices for the pedagogically responsible and ethically grounded integration of AI in special and inclusive education, offering preliminary guidance for educators.

The present study addresses a critical gap in the literature by investigating the use of ChatGPT 4o-mini in the design of Individualized Education Programs (IEPs) for students with learning disabilities (LDs), as implemented by special education teachers in inclusive classrooms in Greece. Adopting a qualitative-dominant design with quantitative support, the study examines the quality of IEP goals created with and without the support of ChatGPT, while also exploring how prior training in IEP development and the amount of planning time may influence outcomes. Additionally, it seeks to capture teacher perspectives on both the opportunities and challenges associated with integrating AI into the IEP development process. The findings aim to inform best practices for the pedagogically responsible use of artificial intelligence in special and inclusive education, offering both empirical insights and practical guidance to the ongoing conversation around AI-enhanced inclusive teaching.

To guide this investigation, the study is structured around five distinct research questions:

Research Question 1. Does the use of ChatGPT by special education teachers lead to a measurable difference in the quality of IEP goals compared to those developed without it?

Research Question 2. Are there observable differences in the content, focus, or domain-specific areas of IEP goals between teachers who utilize ChatGPT and those who do not?

Research Question 3. How does prior training in IEP development and goal writing influence the quality of IEP goals, both with and without the use of ChatGPT?

Research Question 4. Is there a significant difference in the time required for goal writing and overall IEP development between teachers who use ChatGPT and those who do not?

Research Question 5. What are the perceptions of special education teachers regarding the challenges they encounter in designing and developing IEPs for students with LDs, and how do these challenges differ depending on whether ChatGPT is used?

2. Materials and Methods

2.1. Participants

A purposive sampling strategy was employed, combining criterion-based sampling with the snowball technique to recruit participants who were directly relevant to the study's aims. Eligibility was limited to permanent special education teachers working in inclusion classrooms of public primary schools in a city in Northern Greece, where students with formally diagnosed learning disabilities (LDs) were enrolled. All diagnoses had been officially issued by the Center for Interdisciplinary Assessment, Counseling and Support (KEDASY), a state-recognized authority.

The final sample consisted of six special education teachers. Three educators were assigned to the control group and three to the experimental group. All participants were female except for one male in the control group. Their academic and professional back-

grounds varied in terms of formal training and teaching experience, with detailed information provided in Table 1 below.

Table 1. Demographic characteristics of the participants.

Group	Gender	Age Range	PE Code	Academic Qualification	Years in General Ed.	Years in Inclusion	IEP Training
Control (n = 3)	2 Female 1 Male	31–40 (n = 2) 51–60 (n = 1)	PE70 (n = 2) PE71 (n = 1)	BA in General Ed. + MA in SpEd (n = 2) BA in SpEd (n = 1)	7–32 years	4–25 years	Yes (n = 2) 20-h training by KEDASY/University
Experimental (n = 3)	3 Female	31–40	PE71	BA in SpEd (n = 1) MA in SpEd	Not specified	5–12 years	Yes (n = 1) No (n = 2)

Participants in the control group used traditional methods to develop IEPs, while those in the experimental group utilized ChatGPT (version 4o-mini) for IEP planning, with each session lasting approximately 1.5 h. Differences were noted in age, academic qualifications, prior training in IEP development, and years of teaching experience both in general education and in inclusion settings.

2.2. Ethical Considerations

This study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki (1975, revised 2013) and in full compliance with GDPR regulations on data protection. All participants were informed about the study’s aims, procedures, and their right to withdraw at any time without consequence. Written informed consent was obtained from all special educators prior to participation. In addition, parental consent was requested and obtained for the children involved in the educational intervention, ensuring that both the educators and the children they worked with were fully protected.

Given the non-clinical nature of the research, which involved only interviews, questionnaires, and an educational intervention without collection of sensitive biomedical data or invasive procedures, formal Institutional Review Board (IRB) or Ethics Committee approval was not required under national guidelines. Nevertheless, strict measures were applied to guarantee anonymity, confidentiality, and voluntary participation. All data were handled and stored in compliance with GDPR standards, and the use of ChatGPT was closely supervised to ensure responsible application, accuracy, and protection of participant privacy.

2.3. Instruments

The adoption of a qualitative-dominant design with quantitative support enabled both the descriptive assessment of IEP goal quality and in-depth exploration of teachers’ perceptions—an approach aligned with best practices in inclusive education research, where both measurable outcomes and experiential insights are essential (Creswell, 2011). Three primary data collection instruments were employed in this study: a microteaching documentation form, a semi-structured interview protocol, and a self-administered questionnaire.

All participants, regardless of group assignment, completed the microteaching documentation form following each instructional session (lasting 45 min) conducted with a student formally diagnosed with a learning disability (LD), selected by the teacher. This form required the documentation of core instructional components, including the teaching-learning sequence and the student’s current level of performance. These assessments were informed both by the official diagnosis provided by the Center for Interdisciplinary Assessment, Counseling and Support (KEDASY) and by the teacher’s own initial evaluation. Participants were also asked to formulate both long-term and short-term learning objectives across cognitive, social, and metacognitive domains, specifying targeted skills and attitudes.

To promote consistency in goal-setting, all participants were provided with a reference guide based on Bloom's taxonomy.

For the control group, additional data were collected via a self-administered questionnaire developed using Google Forms. This instrument included an introductory section and a main body divided into 9 thematic areas, comprising 43 items aimed at capturing participants' perceptions, practices, and experiences related to IEP development.

In contrast, participants in the experimental group participated in semi-structured interviews, guided by a protocol organized into six thematic sections: (1) introductory background, (2) use of ChatGPT, (3) tool adaptation and effectiveness, (4) support and challenges, (5) perceived advantages and limitations, and (6) future perspectives on AI integration in the IEP design process.

2.4. Procedure

The research procedure commenced with an initial round of telephone contact targeting special education teachers working in inclusion classrooms in a city in Northern Greece who supported students formally diagnosed with learning disabilities (LDs). Eligible participants were randomly assigned to either the experimental or the control group. Both groups received a detailed informational document via email outlining the aims and steps of the study.

For the experimental group, the document contained an introduction to ChatGPT, a concise summary of its potential applications in the context of special education, and an illustrative example of how a teacher might use the tool in designing individualized interventions. It also included guidance on completing the microteaching documentation form, a curated list of action verbs aligned with Bloom's taxonomy, and a set of representative prompts adapted from [Gavriilidou \(2024\)](#) and customized by the research team for use in the special education context. Contact details for technical or pedagogical support were also provided.

Within the experimental group, two participants reported prior experience with ChatGPT. The third, unfamiliar with the tool, received individualized training support from a member of the research team through four face-to-face sessions (1.5 h each). These sessions involved account setup, guided practice, and exploration of prompt design. All experimental participants were given one month to study the materials, engage with the tool, and seek further clarification as needed.

The control group received a structurally identical document, excluding any reference to ChatGPT or AI-related content.

After completing the design and implementation of their microteaching sessions, participants submitted the corresponding documentation forms. Subsequently, those in the experimental group were invited to participate in semi-structured interviews, while those in the control group received an online questionnaire, with a one-week window for submission.

2.5. Quantitative Data Analysis

A total of 368 Individualized Education Program (IEP) goals were developed and analyzed in this study—97 by the control group and 271 by the experimental group. Each goal was evaluated using the Revised IEP/IFSP Goals and Objectives Rating Instrument (R-GORI) ([Notari-Syverson & Shuster, 1995](#)), a validated tool for assessing the quality of IEP goals across four critical dimensions: functionality, generality, instructional context, and measurability.

- Functionality (2 indicators) examines the real-life relevance of the skill targeted by the goal, including its contribution to social participation, daily independence, and

its necessity for task completion. It also considers whether the task would need to be completed by someone else if not mastered by the student.

- Generality (3 indicators) assesses the extent to which the goal can be applied across different environments, people, and materials. It evaluates whether the targeted skill is broadly applicable, transferable, and relevant across disability types and contexts.
- Instructional Context (2 indicators) focuses on whether the skill can be taught within naturalistic settings—such as the classroom or home—by both teachers and caregivers. This dimension also evaluates the goal’s clarity, accessibility, and avoidance of technical language.
- Measurability (3 indicators) concerns the degree to which the goal can be objectively assessed. It requires the specification of observable criteria, including indicators, such as frequency, duration, accuracy, or time constraints, along with clearly defined conditions for performance.

Each of the ten R-GORI indicators was scored dichotomously (0 = absent, 1 = present), yielding a composite score ranging from 0 to 10 per goal. Higher scores represent greater overall quality and alignment with best practices in IEP design.

Statistical analysis of the goal scores was performed using IBM SPSS Statistics (Version 26.0) to examine group differences and relationships between variables. While considerable effort was made to analyze the data rigorously, the very small sample size limits statistical power. As a result, the quantitative findings are context-bound and exploratory, intended to illustrate possible patterns rather than to support statistically generalizable conclusions.

2.6. Qualitative Data Analysis

Following the completion of the semi-structured interviews, all audio recordings were transcribed verbatim and analyzed using thematic analysis, a widely accepted and flexible method in qualitative research (Bryman, 2017). The analysis began with a thorough familiarization with the transcripts, followed by open coding, which produced a total of 249 initial codes.

Next, a combination-based grouping strategy was employed to organize these codes into higher-order categories, also referred to as thematic patterns (Bryman, 2017). This stage involved the annotation, naming, and interpretive classification of emergent themes. Identified themes were subsequently categorized as anticipated, unexpected, difficult to classify, primary, or secondary, in accordance with Creswell’s (2011) framework.

An interpretive approach guided the analysis, allowing for the exploration of both manifest (explicit) and latent (underlying) meanings in the data (Bryman, 2017). Moreover, the researchers examined conceptual relationships and sequential patterns between themes through the development of thematic networks, in line with the procedure outlined by Attride-Stirling (2001). This process enabled a deeper understanding of participants’ perceptions, experiences, and attitudes regarding the use of ChatGPT in the context of IEP development.

3. Results

3.1. First and Second Research Questions

To address the first and second research questions, an analysis was conducted to determine whether the scores assigned to individual and overall IEP goals differed based on the participants’ group membership. An independent samples *t*-test was performed, with the dependent variables being the mean scores for each goal category (as well as total scores), and the independent variable being the group (experimental vs. control).

As shown in Table 2, participants in the experimental group, who used ChatGPT, achieved significantly higher mean scores for short-term skill goals compared to those in

the control group ($t(4) = -2.92, p = 0.043$). However, no statistically significant differences were observed in the other goal categories.

Table 2. Independent samples *t*-test results for goal quality scores by group.

Dependent Variable	Group	M	SD	<i>t</i> (4)	<i>p</i>
Long-Term Cognitive Goals	Control	3.27	2.11	−1.54	0.199
	Experimental	8.03	5.12		
Long-Term Social Goals	Control	1.45	0.78	−1.17	0.328
	Experimental	3.90	2.76		
Long-Term Metacognitive Goals	Control	1.60	1.27	−1.63	0.201
	Experimental	4.70	2.38		
Long-Term Skills	Control	2.37	1.10	−1.37	0.300
	Experimental	9.27	8.79		
Long-Term Attitudes and Dispositions	Control	1.35	4.60	−0.95	0.414
	Experimental	4.63	4.65		
Short-Term Cognitive Goals	Control	3.05	1.91	−1.07	0.362
	Experimental	9.43	7.87		
Short-Term Social Goals	Control	0.50	0.00	−1.38	0.260
	Experimental	1.90	1.35		
Short-Term Metacognitive Goals	Control	3.25	1.77	0.51	0.647
	Experimental	2.67	0.91		
Short-Term Skills	Control	3.20	1.68	−2.92	0.043
	Experimental	6.17	0.51		
Short-Term Attitudes and Dispositions	Control	1.50	0.00	−1.12	0.465
	Experimental	2.95	1.06		
Overall Goal Quality Score	Control	16.80	10.67	−2.10	0.104
	Experimental	52.67	27.60		

Note: M = mean; SD = standard deviation. Statistically significant differences ($p < 0.05$) are bolded.

3.2. Third Research Question

The third research question explored the impact of prior training in IEP development on the overall quality score of IEP goals, while also considering participants' group assignment (experimental vs. control). To investigate this, a multiple regression analysis was conducted with the overall goal score as the dependent variable, and group membership and prior IEP training as independent variables.

As shown in Table 3, prior training in IEP development did not have a statistically significant effect on the overall goal score ($t = 2.08, p = 0.129$). However, teachers in the experimental group achieved significantly higher overall scores on their goals compared to those in the control group ($t = 3.37, p = 0.043$).

Table 3. Multiple Regression Results Predicting Overall Goal Quality Score.

Predictor	B	SE	95% CI	<i>t</i>	<i>p</i>
Constant	−1.77	12.62	[−41.94, 38.41]	−0.14	0.898
Group ¹	45.15	13.39	[2.54, 87.76]	3.37	0.043
Prior IEP Training	27.85	13.39	[−14.76, 70.46]	2.08	0.129

Note: B = unstandardized coefficient; SE = standard error; CI = confidence interval. ¹ Reference group for the "Group" variable is the control group.

3.3. Fourth Research Question

The fourth research question examined whether the time dedicated to the development of microteaching sessions differed between teachers in the experimental and control groups. A chi-square (χ^2) test was conducted to investigate this. It should be noted that the p -value was estimated using Monte Carlo simulation due to the presence of two cells with zero frequencies in the contingency table.

As shown in Table 4, no statistically significant difference was found between the two groups regarding the time spent on microteaching development ($\chi^2 = 3.00$, $p = 0.398$).

Table 4. Chi-square test results for the relationship between microteaching development time and group.

Time Dedicated to Microteaching Development	Control Group	Experimental Group	Total	$\chi^2(2)$	p
Up to one week	1 (33.3%)	3 (100.0%)	4 (66.7%)	3.00	0.398
Two weeks	1 (33.3%)	0 (0.0%)	1 (16.7%)		
Three weeks	1 (33.3%)	0 (0.0%)	1 (16.7%)		
Total	3 (100.0%)	3 (100.0%)	6 (100.0%)		

3.4. Correlation Between Time Investment and Goal Quality

Finally, Spearman's rank-order correlation coefficient was calculated to examine the relationship between the time spent on microteaching development and the mean scores of individual and overall IEP goals.

As shown in Table 5, the time dedicated to microteaching was negatively correlated with the mean score of long-term metacognitive goals ($r = -0.89$, $p < 0.05$) and the mean score of short-term skills ($r = -0.85$, $p < 0.05$). These findings suggest that the more time educators spent preparing their microteaching, the lower their average scores tended to be in long-term metacognitive goals and short-term skill goals.

Table 5. Spearman correlation coefficients between time spent on IEP microteaching development and goal quality scores.

Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Time Spent on Microteaching	1											
2. LT Cognitive Goals	−0.27	1										
3. LT Social Goals	−0.67	0.99 *	1									
4. LT Metacognitive Goals	−0.89 **	0.90 *	0.90 *	1								
5. LT Skills	−0.60	0.41	0.60	0.70	1							
6. LT Attitudes/Dispositions	−0.35	0.90 *	0.80	0.99	0.36	1						
7. ST Cognitive Goals	−0.67	0.99	0.99	0.10	0.60	0.80	1					
8. ST Social Goals	−0.63	0.10	0.10	0.10	0.10	0.40	0.10	1				
9. ST Metacognitive Goals	−0.06	0.10	0.10	0.70	0.10	−0.80	0.10	−0.05	1			
10. ST Skills	−0.85 *	0.43	0.40	0.99	0.40	0.60	0.40	0.46	−0.05	1		
11. ST Attitudes/Dispositions	−0.87	0.50	0.50	0.99	−0.50	0.99	0.50	0.99	−0.50	−0.50	1	
12. Total Goal Quality Score	−0.44	0.94	0.90	0.99	0.49	0.99	0.90 *	0.41	−0.10	0.66	0.99	1

Note: LT = long term; ST = short term. Significant correlations: $p < 0.05$ (*), $p < 0.01$ (**).

3.5. Fifth Research Question: Quantitative Results

Teachers in the control group reported varying experiences regarding the development of Individualized Education Programs (IEPs) using traditional methods. In terms of the initial assessment of students with learning disabilities (LDs), 66.7% described the process as “Moderately” easy, with the time required ranging from one to three weeks, evenly split between the two options (33.3% each).

When asked about the formulation of goals and objectives, 66.7% found the process of writing long-term goals (cognitive, social, and attitudinal) to be “Very” easy. In contrast,

the specification of skills and short-term goals was rated as “Moderately” easy by the majority (66.7%). Responses regarding short-term skills and attitudes were equally distributed across “Not at all,” “Moderately,” and “Very” easy, indicating a lack of consensus among participants.

The selection of instructional themes and planning of learning sequences were described as “Very” easy by 66.7% of participants. In terms of instructional preparation, the majority (66.7%) reported spending 2–3 h daily searching for appropriate materials, while the remaining 33.3% spent 4–5 h. The most frequently used and easily accessible teaching resources included computers, internet-based tools, worksheets, and educational books beyond standard textbooks—all of which were utilized by 100% of the teachers.

Among the teaching methods considered most easy to apply were inquiry-based learning, experiential learning, brainstorming, interdisciplinary approaches, critical thinking strategies, and differentiated instruction. Each was identified by 66.7% of participants as effective and feasible within their teaching contexts.

Regarding assessment practices, 66.7% of the participants reported evaluating instructional effectiveness “Sometimes,” while the remaining 33.3% did so “Always.” Formative assessment was used by 66.7%, whereas 33.3% implemented all three assessment types (diagnostic, formative, and summative). All teachers employed oral and written assessments, along with observational techniques.

With respect to self-assessment, 66.7% engaged in it “Sometimes” and 33.3% “Always.” All participants reported deviating from their initial instructional plans “Sometimes.” When asked about the time invested in developing their IEPs, 66.7% spent over five hours, while 33.3% reported spending one to two hours.

Participants also evaluated the effectiveness and challenges of traditional methods. The aspects of individualization, ease of IEP creation, and monitoring were rated as “Slightly,” “Moderately,” or “Very” effective. Major challenges included difficulty in sourcing appropriate instructional materials (66.7%), insufficient technical support (66.7%), and limited time availability (33.3%).

Finally, 66.7% of teachers rated their IEPs as “Very” effective, while the remaining 33.3% considered them “Moderately” effective ratings that aligned with their perceptions of student progress. In terms of exploring new teaching practices, 66.7% reported doing so “Sometimes,” and 33.3% “Often.” Notably, participants’ intention to use AI tools, such as ChatGPT, as well as their perception of such tools’ usefulness in IEP development, was evenly distributed across the options “Not at all,” “Moderately,” and “Very” (33.3% each), indicating diverse attitudes toward AI integration in special education contexts.

3.6. Qualitative Results

Thematic analysis of the semi-structured interviews revealed six central themes that capture the complex, multifaceted experiences of special education teachers using ChatGPT for the development of Individualized Education Programs (IEPs) for students with learning disabilities (LDs). These themes highlight both the perceived benefits and the limitations of AI integration in inclusive education settings, with implications for practice, policy, and teacher training.

3.6.1. Goal Setting and Goal Attainment

Participants widely recognized ChatGPT’s value in assisting with the formulation of learning goals, particularly in developing well-structured cognitive and long-term objectives. The AI’s capacity to organize and articulate these goals was seen as beneficial for planning and clarity. However, its output for short-term goals was sometimes perceived as overly general or lacking the level of specificity required for immediate instructional use.

P1 explained, *“For long-term goals, it was fine. But for the short-term ones, they came out a bit too general again. Honestly, if I had written them by hand, I would have included more detail, especially in the last two lessons on the ‘ts’ and ‘tz’ sounds.”* Similarly, P2 acknowledged, *“Yes, it helped me a lot with the cognitive goals,”* whereas P3 remained cautious: *“Not particularly, I wouldn’t say so.”*

Regarding goal attainment, most participants reported satisfactory progress toward the learning objectives but underscored the need for follow-up assessments to confirm mastery and skill generalization. As P3 noted, without such reassessment, the appearance of success might mask incomplete or unstable learning gains.

3.6.2. Instructional Support and Material Development

All participants in the experimental group praised ChatGPT as a creative partner in generating teaching materials and lesson ideas. Its ability to suggest multi-sensory, non-traditional resources was particularly valued in meeting diverse learning needs and in supporting differentiated instruction.

P1 described, *“I used modeling clay as ChatGPT suggested, also rice, salt, trachanas. . . we have water beads, images, and the interactive whiteboard. The only difference this time is that I didn’t give out as many worksheets as before because the activities were a bit different.”* For P2, ChatGPT’s influence was not in replacing her planning but in refining it: *“I already had a clear plan in my mind, but now everything felt more structured—ChatGPT had the goals written out, everything was more complete.”*

These accounts suggest that ChatGPT can enrich lesson planning by offering varied instructional options, especially for teachers seeking to move beyond traditional worksheet-based approaches.

3.6.3. Recognized Limitations of ChatGPT

While ChatGPT’s suggestions were generally considered pedagogically sound, participants consistently stressed that the AI’s recommendations were designed for idealized scenarios and did not fully account for the unpredictability of real classroom life.

P2 explained, *“Yes, the methods and practices it suggested were appropriate. But they seemed designed for ideal conditions. In a real classroom, anything can happen—a knock at the door, for example—that you can’t foresee.”* P3 elaborated, *“Teachers must have alternatives in mind and adapt to the student’s mood. Not all days are the same. A well-designed plan might fail if the child is upset or distracted—AI can’t anticipate that.”*

This theme underscores a critical limitation of AI tools: their inability to dynamically respond to real-time emotional, behavioral, or environmental factors without human mediation.

3.6.4. Attitudes Toward ChatGPT

Initial teacher perceptions ranged from skepticism to cautious curiosity. For some, reluctance stemmed from unfamiliarity with the tool or uncertainty about its relevance to their professional context. However, hands-on experience often shifted these attitudes toward greater acceptance and appreciation, especially when the benefits became evident in lesson preparation.

P1 admitted, *“A colleague had suggested using ChatGPT, but I wasn’t really into it. . . I wasn’t in favor, but now I see that it works.”* Such shifts suggest that practical exposure, combined with collegial support, may be essential in fostering teacher buy-in for AI-assisted planning.

3.6.5. Ethical and Pedagogical Concerns

Participants expressed strong reservations about the uncritical or extensive use of AI in education, particularly for younger learners. Concerns centered on the risk of diminishing students’ critical thinking skills and fostering overdependence on automated outputs.

P1 reflected, *“As adults and educators, we can judge whether the information is correct—or at least which part is. But children, I’m not sure they can do that, given their age and stage.”* P2 went further: *“For students in primary or secondary education, I don’t think they should use ChatGPT extensively—it prevents them from thinking. I don’t believe it should even be used under the age of 18 because they don’t know how to use it properly.”*

These comments align with broader ethical discussions in AI in education, emphasizing the need for critical literacy skills alongside technological integration.

3.6.6. The Human Dimension in Teaching

A recurring and emphatic theme was the irreplaceable role of the teacher. While ChatGPT was acknowledged as a useful supplementary tool, it was not viewed as a substitute for the relational, emotional, and adaptive qualities of human teaching.

P2 stated, *“No matter how much AI improves, it can’t replace humans. It lacks consciousness—it works with probabilities. . . a child at such a critical developmental stage needs more than a tool. AI can only go so far—it can’t truly understand, embrace, or support the child.”* P3 echoed this sentiment: *“As a human, you’ll add humor to your teaching, take breaks when you see the student needs it. That’s something only a person can recognize in real-time.”*

This theme reinforces that pedagogical relationships and responsiveness remain beyond the reach of AI technology.

3.6.7. Time Investment and the Need for Training

Two of the three participants reported investing more preparation time when using ChatGPT compared to traditional lesson planning. This additional time was attributed to the learning curve of using a new tool effectively.

P3 proposed a clear solution: *“It would be helpful to have some training or workshops explaining how to use ChatGPT—especially for people who haven’t used it yet. That way we could share ideas and experiences within the school community.”*

Such feedback underscores the importance of targeted professional development to build both technical fluency and pedagogical integration strategies.

In sum, teachers viewed ChatGPT as a supportive and creative aid in IEP development, particularly for structuring goals and generating varied materials. However, they emphasized its supplementary role, noting the irreplaceable human dimensions of teaching, the need for critical oversight, and the importance of targeted training for effective use.

4. Discussion

This study examined the use of ChatGPT in the design of Individualized Education Programs (IEPs) for students with learning disabilities (LDs) by special education teachers in inclusive primary classrooms in Northern Greece. Using a qualitative-dominant design with descriptive quantitative support, the analysis combined thematic insights from teacher interviews with illustrative patterns from R-GORI scores. It also brought attention to persistent issues of digital inequality in access to AI tools and infrastructure—barriers that must be addressed if inclusive education policies are to be translated into equitable and meaningful classroom practices (Pagliara et al., 2024; Adel et al., 2024).

Descriptive R-GORI trends suggested that teachers who used ChatGPT often developed IEP goals with higher quality ratings across several indicators. These tendencies are broadly consistent with the patterns reported by Waterfield et al. (2025) but differ from the clearer advantages noted by Rakap (2023) and Rakap and Balikci (2024). Variations in teacher experience, student characteristics, subject areas, and the specific version of ChatGPT employed may help explain these differences. Given the very small sample size,

these quantitative observations are presented only as illustrative examples and are not intended for generalization.

The qualitative findings provide deeper insight into the opportunities and challenges of AI-assisted IEP development. Teachers valued ChatGPT's ability to support goal setting, structure lessons, and generate diverse, multi-sensory instructional resources, in line with creative applications noted by Baytak (2024), Li et al. (2024), and Rizos et al. (2024). However, participants also noted that AI-generated suggestions often assumed ideal conditions, requiring adaptation to the dynamic realities of inclusive classrooms—a limitation similarly highlighted by Küchemann et al. (2023) and Van den Berg and du Plessis (2023).

Across interviews, teachers emphasized that ChatGPT could not replace the human dimensions of teaching, such as empathy, humor, and real-time responsiveness—points that align with Adel et al.'s (2024) argument that AI should remain a supportive rather than substitutive tool. Ethical concerns also emerged, particularly regarding over-reliance on AI by younger learners, echoing the cautionary perspectives of Lee and Zhai (2024) and Rizos et al. (2024).

Prior training in IEP development did not appear to influence descriptive patterns in goal quality, corroborating Rakap's (2023) finding that training alone may be insufficient without sustained professional development. In some cases, lesson preparation with ChatGPT took longer than with traditional planning methods, likely reflecting the learning curve associated with adopting a new tool. This contrasts with reports by Waterfield et al. (2025) and Yilmaz Can and Durmuş (2024) of time-saving benefits, suggesting that efficiency gains may be more evident once familiarity and integration strategies are established.

Overall, these findings are preliminary and specific to the study's context but contribute to the growing body of literature on AI in special and inclusive education. They suggest that ChatGPT can serve as a valuable planning aid when used critically and adaptively, while reinforcing the importance of teacher agency, targeted professional development, and equitable access to technology. These priorities align with international policy frameworks, including UNESCO's Education 2030 Agenda and the European Agency for Special Needs and Inclusive Education (EASNIE), which advocate for inclusive, equitable, and digitally supported learning environments.

4.1. Limitations and Future Directions

Despite its contributions, this study is subject to several limitations that should be taken into account when interpreting the findings. First, the research was geographically confined to a single urban area in Northern Greece, which may limit the generalizability of the results. Including participants from a broader range of geographic contexts—such as rural areas, island regions, and underserved communities—would offer more comprehensive insights. Second, the very small sample size limited the scope of the analysis and does not allow for statistically generalizable conclusions, and the findings should be regarded as context-bound and exploratory. Future studies involving larger cohorts of special education teachers would allow for more robust conclusions. Third, the study focused exclusively on inclusion classrooms within general education settings. While this focus is valuable, it overlooks the distinct dynamics of specialized educational environments. Expanding the scope to include special education schools and other alternative learning contexts would provide a fuller picture of ChatGPT's utility. Moreover, the student population was limited to those with formal diagnoses of learning disabilities. Although this allowed for targeted exploration, it does not capture the diversity of special educational needs present in inclusive education. Future research should include learners with a wider range of needs—such as ADHD, autism spectrum disorder, and intellectual disabilities—and consider both primary and secondary education settings to broaden applicability.

Another limitation lies in the narrow subject focus, as the intervention concentrated solely on language instruction. Investigating the integration of ChatGPT into other disciplines, including mathematics, history, science, and geography, would enhance understanding of its broader educational relevance. Finally, the study employed the ChatGPT 4o-mini model; however, given the rapid pace of advancements in generative AI, subsequent research should explore the features and implications of newer versions to more accurately assess their evolving potential in special education practice.

4.2. Implications for Practice and Policy

The present study offers important insights into the pedagogical integration of ChatGPT in the development of Individualized Education Programs (IEPs) for students with learning disabilities in inclusive primary education settings. The findings underscore the critical need for structured professional development, not only in the technical use of AI tools but also in the cultivation of critical digital literacy. Teachers identified ChatGPT as a supportive tool for instructional planning yet consistently emphasized that effective integration requires systematic training to empower educators to evaluate, adapt, and meaningfully apply AI-generated content within personalized teaching frameworks.

Equally vital is the development of clear ethical and pedagogical guidelines. While ChatGPT can enhance lesson planning, it cannot substitute the professional judgment, emotional intelligence, and classroom responsiveness of the teacher. As such, educational policies must clearly position AI as a supportive—rather than autonomous—tool. Safeguarding teacher agency, ensuring data privacy, and prioritizing student well-being must remain central to any AI-related initiative in education.

The study also revealed infrastructural disparities, particularly within the control group, where limited access to resources and support systems posed significant challenges. This finding highlights the pressing need for equitable access to digital tools, reliable connectivity, and technical support—especially within special education contexts, where differentiated instruction and tailored interventions are essential for student success.

Moreover, the integration of ChatGPT should be responsive to contextual differences, including school type, student needs, and subject area. While the present study focused on language instruction in inclusive classrooms, future research and policy initiatives should investigate the applicability of generative AI across a broader range of subjects—such as mathematics, science, and social studies—as well as in diverse educational environments, including rural schools and specialized education units.

Considering the rapid evolution of AI technologies, education policy must remain both adaptive and forward-looking. Curriculum frameworks, teacher guidelines, and assessment practices should be regularly reviewed and updated to reflect the changing capabilities of AI, while maintaining a strong focus on educational integrity and ethical safeguards. Ultimately, the responsible and effective use of ChatGPT in special education requires a holistic strategy that combines robust teacher training, adequate infrastructure, clear ethical parameters, and continuous evaluation. Policymakers should consider embedding AI literacy and inclusive digital pedagogy into national teacher education standards, ensuring that all educators are equipped to critically and effectively engage with emerging technologies in diverse and inclusive classroom settings.

Author Contributions: All authors contributed equally to the conception and design of the study, data collection, analysis, interpretation of results, and writing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Due to the non-clinical nature of this study, which involved only interviews, questionnaires, and an educational intervention without the collection of sensitive biomedical data, formal approval by an Institutional Review Board (IRB) or Ethics Committee was not required under national guidelines. All procedures were conducted in accordance with the Declaration of Helsinki (1975, revised 2013) and GDPR regulations.

Informed Consent Statement: Written informed consent was obtained from all participating special educators. In addition, parental consent was obtained for the children involved in the educational intervention.

Data Availability Statement: Data are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Adel, A., Ahsan, A., & Davison, C. (2024). ChatGPT promises and challenges in education: Computational and ethical perspectives. *Education Sciences*, 14(8), 814. [\[CrossRef\]](#)
- Attride-Stirling, J. (2001). Thematic networks: An analytic tool for qualitative research. *Qualitative Research*, 1(3), 385–404. [\[CrossRef\]](#)
- Baytak, A. (2024). The content analysis of the lesson plans created by ChatGPT and Google Gemini. *Research in Social Sciences and Technology*, 9(1), 329–350. [\[CrossRef\]](#)
- Bryman, A. (2017). *Social research methods*. Gutenberg.
- Creswell, J. W. (2011). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Ion.
- Eke, O. D. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, 13, 100060. [\[CrossRef\]](#)
- Gavriilidou, Z. (2024). *Teaching and learning language with ChatGPT*. Kritiki.
- Jung, L. A. (2007). Writing SMART objectives and strategies that fit the ROUTINE. *Teaching Exceptional Children*, 39(4), 54–58. [\[CrossRef\]](#)
- Karaman, M. R., & Göksu, I. (2024). Are lesson plans created by ChatGPT more effective? An experimental study. *International Journal of Technology in Education*, 7(1), 107–127. [\[CrossRef\]](#)
- Khuibut, W., Premthaisong, S., & Chaipidech, P. (2023, December). Integrating ChatGPT into Synectics model to improve high school students' creative writing skill. In *Proceedings of the 31st international conference on computers in education (ICCE 2023)*. Asia-Pacific Society for Computers in Education.
- Küchemann, S., Steinert, S., Revenga, N., Schweinberger, M., Dinc, Y., Avila, K. E., & Kuhn, J. (2023). Can ChatGPT support prospective teachers in physics task development? *Physical Review Physics Education Research*, 19(2), 020128. [\[CrossRef\]](#)
- Lee, G. G., & Zhai, X. (2024). Using ChatGPT for science learning: A study on pre-service teachers' lesson planning. *IEEE Transactions on Learning Technologies*, 17, 1643–1660. [\[CrossRef\]](#)
- Li, Y., Liu, J., & Yang, S. (2024). Is ChatGPT a good middle school teacher? An exploration of its role in instructional design. In *Proceedings of the 3rd international conference on new media development and modernized education (NMDME)*, October 13–15, 2023, Xi'an, China. EAI. [\[CrossRef\]](#)
- Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*, 30(4), 26–29. [\[CrossRef\]](#)
- Notari-Syverson, A. R., & Shuster, S. L. (1995). Putting real-life skills into IEP/IFSPs for infants and young children. *Teaching Exceptional Children*, 27(2), 29–32. [\[CrossRef\]](#)
- Pagliara, S. M., Bonavolontà, G., Pia, M., Falchi, S., Zurru, A. L., Fenu, G., & Mura, A. (2024). The integration of artificial intelligence in inclusive education: A scoping review. *Information*, 15(12), 774. [\[CrossRef\]](#)
- Rakap, S. (2023). Chatting with GPT: Enhancing individualized education program goal development for novice special education teachers. *Journal of Special Education Technology*, 39(3), 339–348. [\[CrossRef\]](#)
- Rakap, S., & Balikci, S. (2024). Enhancing IEP goal development for preschoolers with autism: A preliminary study on ChatGPT integration. *Journal of Autism and Developmental Disorders*, 1–6. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rashed Ibraheem Almohesh, A. (2024). AI application (ChatGPT) and Saudi Arabian primary school students' autonomy in online classes: Exploring students and teachers' perceptions. *International Review of Research in Open and Distributed Learning*, 25(3), 1–18. [\[CrossRef\]](#)
- Rizos, I., Foykas, E., & Georgakopoulos, S. V. (2024). Enhancing mathematics education for students with special educational needs through generative AI: A case study in Greece. *Contemporary Educational Technology*, 16(4), ep535. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rospigliosi, P. A. (2023). Artificial intelligence in teaching and learning: What questions should we ask of ChatGPT? *Interactive Learning Environments*, 31(1), 1–3. [\[CrossRef\]](#)
- Van den Berg, G., & du Plessis, E. (2023). ChatGPT and generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences*, 13(10), 998. [\[CrossRef\]](#)

- Waterfield, D. A., Coleman, O. F., Welker, N. P., Kennedy, M. J., McDonald, S. D., & Cook, B. G. (2025). IEPs in the age of AI: Examining IEP goals written with and without ChatGPT. *Journal of Special Education Technology*. [\[CrossRef\]](#)
- Yilmaz Can, D., & Durmuş, C. (2024, June 18–21). *From AI-generated lesson plans to the real-life classes: Explored by pre-service teachers*. 10th International Conference on Higher Education Advances (HEAd'24), Universitat Politècnica de València, Valencia, Spain. [\[CrossRef\]](#)
- Zemljak, D. (2023). Advanced tools for education: ChatGPT-based learning preparations. *Natural Science Education*, 20(1), 10–19. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.