



A multifaceted evaluation of Master of Arts programs in English language education

^aPınar Yakut ^bMerve Savaşçı

^aÖzyeğin University, Türkiye, pinar.karabacak@ozyegin.edu.tr

^bSakarya University, Türkiye, msavasci@sakarya.edu.tr

APA Citation: Yakut, P., & Savaşçı, M. (2025). A multifaceted evaluation of Master of Arts programs in English language education. *Focus on ELT Journal*, 7(2), 46-64. <https://doi.org/10.14744/felt.7.2.3>

ABSTRACT

Although Master of Arts (M.A.) programs in English Language Education (ELE) are widely offered across Türkiye, there is a notable lack of research that systematically and comparatively evaluates these programs across institutions in terms of their aims, content, implementation processes, and outcomes. In response to this gap, the present descriptive mixed-methods study explored the perspectives of current students and graduates from 23 distinct M.A. programs in ELE in Türkiye, employing the Context, Input, Process, and Product (CIPP) evaluation model. Specifically, the study examined participants' motivations for pursuing an M.A. degree, their evaluations of various program components, and the extent to which these evaluations differed according to individual variables such as gender, teaching experience, and academic achievement. In addition, participants' perceptions of program strengths and weaknesses were investigated, together with their recommendations for program improvement. Data were collected through a questionnaire, open-ended questions, and follow-up interviews. The findings indicated that participants who pursued graduate education with the intention of becoming researchers in academic contexts reported a strong alignment between their professional goals and their program experiences. Moreover, certain subgroups, namely, female participants, more experienced teachers, and students with mid-level academic achievement, tended to evaluate the programs more positively. Participants also identified key strengths of the programs and proposed several suggestions for enhancing program quality, including increasing the range of elective courses, updating and enriching course content, and fostering greater curricular diversity. Overall, the findings provide valuable implications for program developers, faculty members, and prospective M.A. candidates, contributing to the ongoing improvement of graduate education in English Language Education.

Keywords

Program evaluation, master's program evaluation, M.A. program, English language education, CIPP.

Article History

Received	:	05.07.2025
Revised	:	22.11.2025
Accepted	:	24.12.2025
Published	:	30.12.2025

Type

Research Article

Introduction

Teacher education is a continuous process that begins with undergraduate initial teacher education (ITE) programs and continues through various in-service professional development (PD) activities throughout teachers' careers. Continuing PD may also involve activities

⁷**Corresponding Author:** pinar.karabacak@ozyegin.edu.tr

Copyright for this article is retained by the author(s), with first publication rights granted to Focus on ELT Journal.

provided by external expertise, such as graduate education programs. Graduate education, which builds upon undergraduate studies, encompasses Master of Arts (M.A.) and Doctor of Philosophy (Ph.D.) programs, providing specialization in the field, with M.A. programs representing the initial stage of this advanced academic trajectory. This study focuses on the systematic evaluation, which refers to “the process of determining the merit, worth, and value of things” (Scriven, 1991, p. 1), of M.A. programs in English Language Education, with the aim of assessing their quality and guiding continuous improvement efforts.

In the context of Turkish graduate education, M.A. programs are designed to cultivate field specialists who can engage with and interpret knowledge through scientific research (Council of Higher Education-CoHE, 2020). These programs are subject to systematic evaluation to maintain teaching standards, ensure program quality, as well as foster and guide continuous development. Program evaluation involves gathering and examining essential information to enhance a program and assess its effectiveness (Brown, 1995). Similarly, Robinson (2003) defines program evaluation as “the collection, analysis, and interpretation of information ...for forming judgments about the value of a particular program” (p. 199, as cited in Peacock, 2009, p. 261). Given its broad implications for educational programs, educators, students, and stakeholders, program evaluation serves as a critical mechanism that contributes to curriculum development, informs institutional decision-making, and identifies program strengths and weaknesses, ultimately supporting continuous program improvement.

Accordingly, this study adopted the Context, Input, Process, and Product (CIPP) evaluation model to investigate the perspectives of students and graduates on Master of Arts (M.A.) programs in English Language Education (ELE). Although various evaluation models are available (e.g., Brown, 1995; Fitzpatrick et al., 2010; Stufflebeam, 1983), the most widely used is Stufflebeam’s (1983) CIPP model. This framework provides a comprehensive approach for assessing program effectiveness, examining learning environment and objectives (*Context*), resources and strategies (*Input*), implementation processes (*Process*), and outcomes or results (*Product*). Adopting the CIPP model in this study enabled a comprehensive and holistic assessment of M.A. programs in ELE, providing insights into their strengths, areas in need of improvement, and overall effectiveness.

Literature

Several empirical studies have evaluated English Language Education (ELE) educational programs. Nonetheless, most of the earlier research within the Turkish context appears to focus mainly on the evaluation of undergraduate programs (e.g., Okumuş & Daloğlu, 2024) or English preparatory programs (EPPs) (e.g., Atar et al., 2020; Bayram & Canaran, 2019; Kuzu et al., 2021; Mede, 2012). Comparatively, earlier investigations on graduate program evaluation reveal a considerably more limited body of literature—a critical observation highlighted by some scholars (e.g., Çıtak, 2021; Döner-Arday, 2022; Kırmızı, 2011; Öztürk, 2015).

Among studies specifically targeting the evaluation of graduate programs, some were conducted primarily at the Ph.D. level (e.g., Bilican, 2014; Küçükoğlu, 2015). According to those focusing on M.A. program evaluation, although data from international contexts are

limited, emerging evidence from Iran suggests the need for a detailed evaluation. In a mixed-methods comprehensive study, Foroozandeh et al. (2007) evaluated the Teaching English as a Foreign Language (TEFL) curriculum in M.A. programs at nine universities in Iran, using the CIPP model. Data were collected from students enrolled in or graduated from an M.A. program, instructors, and administrators. Regarding program strengths, participants highlighted the presence of dedicated instructors who are knowledgeable and up to date in their field. However, the findings also indicated the necessity for revisions in both the curriculum and delivery, specifically in the program objectives, course design, and course distribution. Other weaknesses comprised the program implementation method and the misalignment between the program objectives and the courses.

In the Turkish graduate education context, comparatively more studies evaluating M.A. programs are available. Most of these studies focused on the specific contexts of individual universities. For example, Kesli-Dollar et al. (2014) evaluated the MA-TEFL Program at a foundation university. Students, academic staff, administrators, and graduates participated in a survey and were interviewed. Most students were highly or moderately satisfied with the courses, instructors, program, assessments, and course materials, with satisfaction rates ranging from the lowest to the highest, respectively. Some thought the program was demanding, making them feel overwhelmed, anxious, and insufficient. Program strengths included qualified academic staff, practical and research-oriented assignments, well-structured syllabi, comprehensive course materials, and a diverse range of courses. Conversely, weaknesses included limited resources to support students with tasks and assignments, excessive workload, and unnecessary theoretical courses. A similar but more extensive study was conducted by Öztürk (2015), who evaluated the M.A. program offered at a foundation university. The study elaborated upon students' motivations for pursuing an M.A. degree and their perceptions of the program's content, instruction, resources, and anticipated outcomes. As reported, participants were delighted with the program, and the instructors were highly acknowledged among the program's strengths; nevertheless, minor adjustments were suggested for the program content.

Another study, employing M.A. program students, was conducted by Çıtak (2021) at a public university in Türkiye. Ten students who had completed their courses evaluated the program through a questionnaire and interviews. Their evaluation embraced their motivational reasons for enrollment, experiences, program strengths, and shortcomings. Strengths concentrated upon qualified, experienced, and accessible academic staff; good rapport between students and academic staff; research-focused courses; and course variety. At the same time, some shortcomings were noted, including the limited use of technology and the use of outdated instructional methods and course materials. In a comparatively recent study, Döner-Arday (2022) investigated the perceptions of graduates and students regarding an M.A. program in English Language Teaching (ELT) at a leading public university. The focus was on program components, challenges, and needs. Analysis of the questionnaire and interviews suggested that participants primarily pursued M.A. programs due to their aspiration to pursue an academic career and later obtain a Ph.D. They considered the academic staff to be supportive and of high quality, the program to be up-to-date and research-oriented, and there to be a variety of courses

from different fields. They also identified some program weaknesses, including high expectations from students, limited collaboration opportunities, insufficient course feedback, overcrowded classrooms, excessive student workload demands, inconvenient class schedules and timings, limited technology integration, and a limited variety of course offerings per semester.

Compared to these, only two comprehensive studies (i.e., Kırmızı, 2011; Zeybek & Tunçer, 2020) have evaluated M.A. programs across different universities. One such study was conducted by Kırmızı (2011), who undertook a mixed-method study evaluating M.A. programs offered at 16 universities involving students, graduates, and instructors. Students' desire to continue their academic pursuits was the key driver behind their decision to pursue an M.A. degree. Overall, moderate program satisfaction levels were reported; however, the programs needed more advanced technology and up-to-date courses. The other study was conducted by Zeybek and Tunçer (2020), who sampled graduate students (four M.A. and four Ph.D. students in English Language Teaching, ELT) at different state universities. Although researchers did not solely focus on M.A. students or specify the number of universities included in the study, they investigated graduate students' feelings and problems in the Turkish setting, adopting a phenomenological qualitative approach. A mixed range of both positive and negative feelings was found, including anxiety, confusion, (dis)satisfaction, (de)motivation, (un)confidence, and feeling under pressure. Some suggested increasing access opportunities to international databases and statistics courses. In contrast, some voiced the incongruence between their initial expectations and reality, highlighting their dissatisfaction with M.A. courses, which they attributed to a lack of sophistication.

Overall, the review of literature suggests that only a few studies have evaluated M.A. programs in ELE in Türkiye. In other words, few studies specifically focus on M.A. programs, and those that do often concentrate on programs within individual universities. That is, the studies in the literature seem to be smaller in number and narrower in scope regarding graduate program evaluation. By conducting a comprehensive study on the perspectives of students and graduates on M.A. programs in ELE in the Turkish setting, the current study aims to contribute to the field.

Purpose of the Study and Research Questions

This study aims to evaluate M.A. programs in English Language Education within the context of Turkish graduate education. Specifically, it examines the aims and motivations of students and graduates pursuing an M.A. degree, as well as their overall satisfaction with the program in which they are enrolled or have graduated. With these aims in mind, the following research questions guided this study:

- 1) With what purpose and motivation did students and graduates in Türkiye start the Master of Arts (M.A.) program in English Language Education?
- 2) Do master's students' and graduates' M.A. program evaluations vary according to a) gender, b) teacher seniority, and c) undergraduate academic achievement?
- 3) What are the opinions of master's students and graduates regarding the M.A. program they are enrolled in/ graduated from concerning a) content (purpose and objectives), b) teaching, c) resources, and d) program achievements?
 - 3.1) What are the strengths and weaknesses of the M.A. program they enrolled in/ graduated from, as perceived by them?
 - 3.2) What are the suggestions for the improvement and development of the graduate program they are enrolled in/graduated from, as perceived by them?

Methodology

Design

Since the study aimed to investigate the perspectives of students and graduates on M.A. programs in English Language Education, a descriptive research design was adopted. It has been posited that the descriptive approach enables the investigation of the target variables utilizing various research techniques (Malagón-Amor et al., 2015). Following the aims, a mixed-method methodology was adopted, ensuring the collection, analysis, and combination of quantitative and qualitative data (Creswell, 2012). Mixed-method studies provide “a more complete understanding of research problems than does the use of either approach alone” (Fraenkel et al., 2022, p. 503).

Setting and Participants

Participants were students and graduates from various M.A. programs (with thesis) in English Language Education (ELE), also known as English Language Teaching (ELT), in Türkiye. In the Turkish context, M.A. programs typically span two years and aim to prepare field experts who can conduct scientific research, interpret findings, and apply evidence-based knowledge to their professional practice. According to the CoHE's (2020) framework regulation, M.A. programs comprise at least eight courses, including a seminar course and a thesis project, totaling 120 ECTS credits, provided that one academic term includes no fewer than 60 ECTS credits. After getting approval from the university's Research Ethics Committee for this study, invitations were sent to all universities (offering M.A. programs in ELE and ELT only) nationwide, to which students and graduates from 23 different programs responded. Participation was voluntary, where participants provided their informed consent before responding to the instruments. The study comprised 69 participants (49 female, 19 male, and one who preferred not to disclose their gender), whose ages ranged from 21 to 54 ($M = 29.93$ years, $SD = 6.057$ years). The sample included M.A. program students and graduates, whose demographics are presented in Table 1 below. This table summarizes that the majority of participants were female, employed at universities' English preparatory schools, and held undergraduate degrees in English Language Teaching.

Table 1. Demographic Characteristics of the Participants

Variables	Category	Frequency (f)	Percentage (%)
Gender	Female	49	71
	Male	19	27.5
	Prefer not to say	1	1.4
Current status in the M.A. program	Taking M.A. courses	17	24.6
	Writing an M.A. thesis	21	30.4
	M.A. program graduate	31	44.9
Field of Employment	Kindergarten	3	4.3
	Primary school	5	7.2
	Secondary school	13	18.8
	High school	11	15.9
	University (at English preparatory school)	21	30.4
	University (at faculty)	4	5.8
	Other	12	17.4
Teacher seniority	Novice (0-5 years)	37	53.6
	Mid-career (6-10 years)	21	30.4
	Experienced (11+ years)	11	15.9
Undergraduate degree in	English Language Teaching	57	82.6
	Translation and Interpretation	1	1.4
	English Culture and Literature	1	1.4
	American Culture and Literature	1	1.4
	English Language and Literature	9	13

Data Collection and Analysis

Three different instruments were employed to promote validity and triangulate findings: a questionnaire, open-ended questions, and semi-structured interviews. The questionnaire was partially adapted from Kırılmaz (2011) with his permission, obtained via email. The original questionnaire consisted of five sections; however, two sections (*Evaluation of Courses* and *Program Goals*) were excluded as they were deemed irrelevant to the current study's aims. The adapted questionnaire, therefore, consisted of three sections: (1) demographic information (12 items), (2) aims and motivation factors for pursuing an M.A. degree (5 items), and (3) program evaluation scale (52 items), totaling 69 items. To support quantitative data, open-ended questionnaire items (see Appendix A) and interview questions (see Appendix B) were

developed based on the study objectives and relevant literature, ensuring alignment with the CIPP evaluation model. The questions were developed by the researchers in consultation with two experts with doctoral degrees (one from the ELE and the other from Educational Sciences) to ensure clarity, relevance, and comprehensiveness. A pilot study was conducted to refine the questions before data collection. All interviews were audio-recorded with participants' consent and transcribed verbatim. The transcriptions were carefully reviewed against the recordings to ensure accuracy, providing a reliable dataset for subsequent qualitative analysis.

Ethical approval for this study was obtained from the Sakarya University Ethics Committee (Document No: E-61923333-050.99-105433, dated February 10, 2022) prior to data collection. Following approval, all universities offering M.A. programs in ELE were formally invited to participate, and the online instrument set was distributed to potential participants through Google Forms. Informed consent was obtained electronically from all participants before their access to the questionnaire. Participants were informed about the voluntary nature of the study, the principles of anonymity and confidentiality, and they confirmed their consent by selecting the required approval statement on the first page of the online instrument. After administering the questionnaire and open-ended items, semi-structured one-on-one interviews were conducted with ten participants, which were audio-recorded and transcribed with the interviewees' informed consent, which was obtained orally and recorded on video. A maximal variation sampling strategy was used to determine the interviewees, selecting those who demonstrated "a diversity of perspectives or characteristics" (Fraenkel et al., 2022, p. 391). Different satisfaction levels, university type (i.e., state or foundation), and current status (i.e., student taking courses, students writing their theses, or graduate students) guided interviewee selection. For analyzing quantitative data, descriptive statistics, frequencies, and non-parametric inferential statistics tests (i.e., the Shapiro-Wilk test, the Mann-Whitney U test, and the Kruskal-Wallis test) were performed using SPSS 26.0. Normality tests indicated that the data were not normally distributed; therefore, non-parametric tests were conducted for inferential analysis.

Meanwhile, qualitative data were analyzed using inductive content analysis. Following the guidelines proposed by Saldaña (2011), the researchers conducted a detailed and iterative examination of the interview transcripts. Each transcript was read and re-read multiple times to gain a thorough understanding of the data. Key concepts were systematically identified and then organized into meaningful categories. These categories were further analyzed and abstracted to develop overarching themes (and codes) that reflect patterns and insights across the dataset. This comprehensive, step-by-step approach enhanced both the credibility and the transparency of the qualitative analysis, ensuring that interpretations were firmly grounded in participants' responses.

Results

Quantitative Data

The results concerning the perspectives of graduates and students on the M.A. programs are illustrated as follows. Research Question 1 (RQ1) focused on participants' purposes and

motivations for starting an M.A. program in English Language Education, and the quantitative results also included participants' descriptions of their anticipated or preferred professional career paths, as presented in Table 2.

Table 2. Professional career choices in pursuing M.A. studies

Professional career choice	M	SD	1= The least important	2= Not important	3= Important	4= Very important
Researcher in an academic setting	3.42	.77	2.9%	8.7%	31.9%	56.5%
English teacher in a state school	2.37	1.04	26.1%	26.1%	21.9%	15.9%
English teacher in a private school	1.92	.98	43.5%	29%	18.8%	8.7%
Management or administration	1.84	.88	42%	37.7%	12.5%	5.8%

As shown in Table 2, most participants prioritized becoming researchers in academic settings ($M = 3.42$, $SD = .77$), indicating a strong inclination toward academic and research-oriented careers. In contrast, relatively low mean scores were reported for teaching positions in state or private schools, as well as for administrative roles, suggesting that these career paths were viewed as less appealing. Overall, the results suggest that participants' professional aspirations were primarily focused on pursuing advanced academic qualifications and contributing to research in the field of ELT. Participants were also asked to rate the influential factor(s) that motivated them to start their M.A. studies, as presented in Table 3.

Table 3. Influential determinants to start an M.A. study

Factors	M	SD	1= The least important	2= Not important	3= Important	4= Very important
Personal intellectual enrichment	3.6	.64	1.4%	4.3%	26.1%	68.1%
Advanced degree required for career advancement	3.47	.83	4.3%	8.7%	21.7%	65.2%
Primary career choice	3.44	.73	1.4%	10.1%	30.4%	58%
Increased income-earning potential	2.79	.97	13%	20.3%	40.6%	26.1%
Change of career	2.66	1.12	21.7%	18.8%	30.4%	29%

As shown in Table 3, the main determinant influencing participants' decisions to pursue an M.A. degree was personal intellectual enrichment ($M = 3.6$, $SD = .64$), followed by the requirement of an advanced degree for career advancement. These results suggest that participants were primarily motivated by academic growth and the desire to enhance their professional qualifications rather than financial or career change motives. Overall, the M.A.

programs were viewed as a meaningful step toward both intellectual and professional development.

To offer a concise and coherent overview of participants' evaluations of the M.A. programs, the descriptive statistics of the program evaluation scale were synthesized at the sub-category level. As each sub-category comprised a substantial number of items, item-level findings were consolidated to enhance clarity and readability. Accordingly, Table 4 summarizes the ranges of mean scores and standard deviations for each sub-category, providing an integrated view of participants' perceptions across the main components of the program.

Table 4. *Descriptive statistics of program evaluation sub-categories*

<i>Sub-categories</i>	<i>Number of items</i>	<i>M</i>	<i>SD</i>
Program description	9	2.50 – 4.37	0.84 – 1.30
Departmental support	3	3.20 – 4.10	0.97 – 1.14
Atmosphere in the department	4	4.07 – 4.28	0.94 – 1.17
Program instruction and evaluation methods	9	3.82 – 4.42	0.81 – 1.17
Classroom management and cooperation skills	3	3.15 – 3.60	1.26 – 1.39
Program resources	4	3.31 – 3.97	0.99 – 1.34
Program content	12	3.49 – 4.30	0.86 – 1.27
Overall program evaluation	7	3.95 – 4.39	0.82 – 1.05

Note. Values indicate the range of mean scores and standard deviations across items within each sub-category.

Overall, the findings indicate that participants reported generally positive perceptions across most program evaluation sub-categories, suggesting a high level of satisfaction with both the academic and institutional aspects of the programs. Regarding *program description*, the mean range ($M = 2.50-4.37$) indicates that participants largely perceived the programs as professionally organized and respectful toward M.A. candidates. While high mean scores on several items reflect positive views of faculty treatment and program structure, the observed variation suggests that certain administrative or organizational aspects were experienced differently across programs. In terms of *departmental support*, the mean range ($M = 3.20-4.10$) suggests that participants generally evaluated the support provided by faculty and departments positively. Most participants agreed that academic staff were helpful and that the programs contributed to their preparation for future professional work; however, comparatively lower ratings related to employment support indicate an area where departments may enhance their guidance for graduates. Participants' evaluations of the *atmosphere in the department* were consistently high ($M = 4.07-4.28$), indicating a supportive academic climate characterized by mutual respect, cooperation, and effective communication between faculty members and M.A.

candidates. These findings underscore the importance of positive interpersonal relationships and a humane departmental environment in shaping the experiences of graduate students.

With respect to *program instruction and evaluation methods*, the results reveal high levels of satisfaction ($M = 3.82-4.42$). Participants perceived the instructional practices as effective in promoting intellectual development, reflective teaching, and research competence. Although overall evaluations were positive, relatively lower scores for course coherence suggest that stronger alignment across courses could further enhance instructional quality. Regarding *classroom management and cooperation skills*, the mean range ($M = 3.15-3.60$) was comparatively lower than that of other sub-categories. This finding suggests that while programs adequately support theoretical and research-oriented development, they may place less emphasis on practice-oriented components related to classroom management, highlighting a potential area for improvement. Participants' evaluations of *program resources* yielded moderate mean scores ($M = 3.31-3.97$), indicating general satisfaction with institutional resources, particularly library holdings. Nevertheless, lower evaluations of specialized facilities and technological support suggest that improvements in material and infrastructural resources could further support teaching and research activities.

Finally, the results for *overall program evaluation* indicate generally positive perceptions of the M.A. programs ($M = 3.95-4.39$). Participants reported high levels of satisfaction with their overall learning experiences and perceived the programs as valuable for their academic and professional development. Despite these positive evaluations, the variability across sub-categories suggests that continued efforts to address identified weaknesses may further enhance program quality.

Comparative analyses were conducted to explore potential differences in participants' overall program evaluations according to gender, teacher seniority, and undergraduate academic achievement (RQ2). The results are summarized in Table 5 below.

Table 5. *Program evaluation according to demographic variables*

Variable	Group	M	SD	Test	p
Gender	Female ($n = 49$)	206.67	35.73	Mann–Whitney U	.248
	Male ($n = 19$)	193.21	39.42		
	Prefer not to say ($n = 1$)	216.00	–		
Teacher seniority	Novice (0–5 years) ($n = 37$)	197.70	39.55	Kruskal–Wallis	.517
	Mid-career (6–10 years) ($n = 21$)	207.19	37.20		
	Experienced (11+ years) ($n = 11$)	213.45	23.08		
Academic achievement	Low ($n = 14$)	185.57	44.60	Kruskal–Wallis	.193
	Mid ($n = 31$)	210.29	33.21		
	High ($n = 24$)	204.04	34.29		

Note. Mean scores represent total program evaluation scores. No statistically significant differences were found across demographic variables.

With respect to gender, female participants ($n = 49$) reported a higher overall evaluation score ($M = 206.67$, $SD = 35.73$) than male participants ($n = 19$; $M = 193.21$, $SD = 39.42$). However, the Mann–Whitney U test results demonstrated that this difference was not statistically significant ($U = 381$, $p = .248$). This suggests that, although females tended to

evaluate the programs more favorably in terms of mean scores, gender did not function as a determining factor influencing participants' overall program evaluations. The single response in the "prefer not to say" category ($n = 1$) yielded a high score ($M = 216.00$); nevertheless, this value should be interpreted cautiously due to the very small group size and its limited representativeness. Regarding teacher seniority, novice teachers (0–5 years; $n = 37$) reported a lower mean score ($M = 197.70$, $SD = 39.55$) compared to mid-career teachers (6–10 years; $n = 21$; $M = 207.19$, $SD = 37.20$) and experienced teachers (11+ years; $n = 11$; $M = 213.45$, $SD = 23.08$). Despite this gradual increase in mean scores across seniority levels, the Kruskal–Wallis H test results indicated that the differences among the three groups were not statistically significant ($\chi^2(2) = 1.319$, $p = .517$). These findings suggest that participants' evaluations remained largely consistent, regardless of their teaching experience. However, the higher evaluations among more experienced teachers may reflect a slightly stronger alignment between program content/outcomes and the needs of teachers at later career stages. Finally, in terms of undergraduate academic achievement, mid-achieving participants ($n = 31$) reported the highest mean score ($M = 210.29$, $SD = 33.21$), followed by high-achieving participants ($n = 24$; $M = 204.04$, $SD = 34.29$), while low-achieving participants ($n = 14$) reported comparatively lower evaluations ($M = 185.57$, $SD = 44.60$). Nevertheless, the Kruskal–Wallis H test results revealed that these group differences were not statistically significant ($\chi^2(2) = 3.287$, $p = .193$). Accordingly, participants' overall evaluations of the M.A. programs did not vary significantly based on their undergraduate academic achievement, indicating that satisfaction with the program was not strongly associated with prior academic performance.

Taken together, the findings summarized in Table 5 demonstrate that overall program evaluations did not differ significantly according to gender, teacher seniority, or undergraduate academic achievement. Although certain subgroups (female participants, more experienced teachers, and mid-achieving participants) tended to report higher mean scores, these tendencies did not translate into statistically significant differences. This pattern suggests that participants perceived the quality and effectiveness of the M.A. programs in relatively similar ways across demographic categories, implying that program experiences and satisfaction levels may be shaped more by program-level factors (e.g., course content, supervision quality, and resources) than by individual demographic characteristics.

Qualitative Data

The qualitative data, collected through open-ended questions and semi-structured interviews, also provided deeper insight into participants' perceptions of the M.A. program, highlighting both its strengths and areas for improvement.

As shown in Table 6, the most frequently mentioned strengths were related to instructors and academic staff, followed by course content and overall program content. Participants emphasized the support and guidance provided by supervisors, as well as the instructors' knowledge and expertise, suggesting that faculty quality is a central factor contributing to positive student experiences ($f = 27$ for instructors; $f = 22$ for qualified and knowledgeable staff).

Table 6. Program strengths

Themes	Frequency (f)	Codes	Frequency (f)
Instructors /Academic Staff	27	Supervisor's support & supervision Qualified & Knowledgeable	22 5
Courses	17	Up to date Content of the courses	6 11
Content	15	Satisfactory training in research skills Content of the courses	3 12

Regarding courses and content, many participants noted that the curricula were generally up-to-date and incorporated contemporary methodologies in the field. The emphasis on research skills, reflected in the participants' responses ($f = 12$), indicates that programs successfully foster both theoretical understanding and practical competencies, preparing students for academic and professional development. These findings suggest that while programs are perceived as relevant and satisfactory, the quality of instruction and the alignment of course content with current trends are particularly pivotal in shaping students' overall satisfaction. The following excerpts exemplify participants' perspectives:

Some professors are really qualified in both teaching and researching. They help us to be good researchers and a good teacher at the same time. So, what they teach, regardless of the course, becomes valuable. (P66)

I was mostly satisfied with the faculty in terms of their knowledge about the field. The content of the MA program was satisfactory, though there might have been a variety of elective courses to choose from. (P50)

Certain courses and instructors, such as SLA, Current trends in ELT, and Research basics, were positively influential on grasping the requirements of field knowledge and conducting research in our field. (P9)

Content / Because I had a chance to explore different areas in ELT. (P39)

Content of the lessons. They are designed to teach us how to do research and collect and analyze data. I have found it very satisfactory. (P58)

In my opinion, the courses generally were good and met our needs as master's students. (P42)

Beyond mere satisfaction, these qualitative findings reveal that the effectiveness of programs is closely tied to faculty expertise, the quality of supervision, and the relevance of course content to professional and research needs. This underscores the importance of maintaining high instructional standards and continuously updating curricula to ensure that programs not only meet students' expectations but also support their long-term academic and career growth.

The qualitative data also highlighted several areas of concern in M.A. programs, as summarized in Table 7. The most frequently cited weakness was inadequate instructor feedback ($f = 21$), indicating that participants felt they received limited guidance and evaluation from their supervisors, which may have hindered their academic development and confidence. Course-related issues were also noted, including the repetitive nature of some courses and the lack of elective options ($f = 16$), suggesting that the curricula may not fully address students' diverse interests or provide sufficient opportunities for specialization. Finally, resource

limitations were mentioned ($f = 9$), such as technological shortcomings and library insufficiencies, which could restrict students' access to up-to-date materials and learning tools.

Table 7. Program weaknesses

Themes	Frequency (f)	Codes	Frequency (f)
Feedback	21	Inadequate feedback	21
Courses	16	Inadequate elective courses	6
		Needs to be updated	10
Resources	9	Technological shortcomings	5
		Library	4

Regarding feedback as a program weakness, a participant stated the following:

The feedback we have been receiving and the way courses are conducted. (P16)

This highlights a perceived insufficiency in feedback and concerns about course delivery, suggesting that programs could benefit from more structured and constructive feedback mechanisms, as well as improvements in instructional methods. Some also referred to the necessity for course updates, highlighting the limited number and variety of elective courses. For instance, one interviewee indicated the following:

There should be more options to select courses according to the needs and field of interest of MA students. (P33)

Several participants also referred to their dissatisfaction regarding the limited M.A. program resources as follows:

There are some problems about the resources of the institution, and technological shortcomings that need to be addressed. Though minor, it causes some dissatisfaction. (P56)

Maybe resources because some of them were not available, thus, we had to pay for it. (P40)

As per the suggestions, they concentrated on content, courses, and instructor feedback. Some excerpts are provided as follows:

Feedback sessions should be more while conducting research. (P47)

I think the courses should be altered because most of the courses are the same as the courses we took in the undergraduate program, such as Language Testing, Research Methods). (P62)

I didn't like the courses focusing on old-fashioned models in ELT. There could be more innovative courses. It would be more effective if the theoretical background of the topics covered were touched upon more. (P14)

More elective courses could be provided for different needs and interests such as Neurolinguistics, and Psycholinguistics. ELT is not all about methods and techniques. (P64)

Beyond identifying these weaknesses, the findings suggest that the effectiveness of programs could be strengthened by improving the quality and frequency of instructor feedback, diversifying course offerings, and enhancing the availability of resources. Addressing these

areas is essential not only for increasing student satisfaction but also for fostering a more robust academic and professional training environment that aligns with current standards in the field.

Discussion

The first research question (RQ1) investigated the aims and motivations of M.A. students and graduates, and results revealed that personal intellectual enrichment was the primary driving force. Regarding professional career choices, participants mostly favored working as academics, and pursuing a Ph.D. degree was another career path they desired in their academic careers. Çıtak (2021), Döner-Arday (2022), Kesli-Dollar et al. (2014), and Kırmızı (2011) documented congruent findings, noting that participants predominantly aspired to be academics and pursue Ph.D. degrees in the future. Participants likewise aspired to teach adult groups and work as instructors at universities, according to the results. Additionally, the majority of participants identified personal intellectual enrichment as the most crucial factor and influential predictor when they began their M.A. studies. In other words, M.A. students and graduates viewed M.A. programs as a fundamental step in pursuing further studies as well as a means of advancing personally and intellectually. These findings align with those of Küçüköğlu (2015), who also reported that the majority of Ph.D. students preferred to work as academics at universities rather than as English language teachers in other educational contexts. Similarly, participants in Kırmızı's (2011) study aspired to work as researchers in an academic setting. The students enrolled in the program sought to enhance their intellectual and professional development, which may suggest that M.A. programs are considered a crucial step in post-graduate education. The opportunity to collaborate with professors and the program's reputation were other noteworthy factors that affected students' decisions to enroll in a particular M.A. program. In a similar vein, Deniz-Yücer (2020) indicated that the opportunity to collaborate with a specific faculty member and the program's reputation were the two primary inducements for enrolling. Taken together, these findings strengthen the *Context* component of the CIPP model, indicating that students' motives, expectations, and long-term academic goals are largely aligned with what the programs intend to offer. In conclusion, findings of earlier studies (i.e., Döner-Arday, 2022; Kırmızı, 2011; Küçüköğlu, 2015; Öztürk, 2015; Deniz-Yücer, 2020) were consistent with those of the current study within the scope of the first research question.

Research Question 2 (RQ2) aimed to investigate whether evaluations of master's students and graduates in their M.A. programs differ according to their gender, teacher seniority, and undergraduate academic achievement. Findings did not show a significant difference in these variables. However, concerning gender, females' evaluation mean scores were higher than those of males, which is in line with previous studies (i.e., Kırmızı, 2011; Küçüköğlu, 2015). According to teacher seniority, the mean scores of the participants varied, with experienced teachers receiving more favorable evaluations and higher mean scores; nevertheless, this difference was not statistically significant. In a similar vein, Kırmızı (2011) reached the same conclusion, stating that graduates with more than 11 years of teaching experience possessed better knowledge of the goals of an M.A. program compared to those with less experience. Indeed, Küçüköğlu (2015) emphasized in her Ph.D. program evaluation study

that those with less teaching experience often exhibit greater enthusiasm for advancing their professional skills. Regarding academic achievement, although mid-achieving participants had the highest mean scores for the M.A. program evaluation, no statistically significant differences were found between the groups. Overall, the study found that female participants, experienced teachers, and mid-achievers had more favorable thoughts and perspectives regarding the appraisal and evaluation of the M.A. programs; however, there was no statistically significant difference among these subgroups. This pattern may indicate that while demographic variables are not strong determinants of program satisfaction at a statistical level, they still offer insights into specific learner profiles, contributing to the *Input* and *Process* components of the CIPP model, particularly regarding how different groups engage with and perceive program elements.

In line with Research Question 3 (RQ3), which sought to understand the perceptions of master's students and graduates regarding the M.A. programs they are enrolled in or have graduated from, descriptive analysis revealed that most participants were satisfied with the program descriptions of their M.A. programs. Participants felt appreciated, faculty members had positive attitudes toward them, and programs made a concerted effort to hire qualified instructors to meet students' needs, echoing Döner-Arday's (2022) findings. Although most participants were moderately satisfied with the program resources and the adapted language teaching materials, the findings indicated that overall satisfaction was not very high. It was suggested that the resources might be enhanced, and teaching/adapting foreign language teaching materials could be enriched in courses identified as highly important by many scholars in the field (Garton & Graves, 2014; Tomlinson, 2023). All in all, the program evaluations of the majority of participants were significantly high, and the subcategories in which they were most satisfied were departmental support, the atmosphere provided in the department, and the overall evaluation of the M.A. programs. These perceptions align closely with the *Process* and *Product* components of the CIPP model. Participants valued the instructional atmosphere and the quality of the instructor, while also acknowledging that program outcomes, such as academic development, were largely positive.

Within the scope of Research Question 3.1. (RQ3.1.), elaborating on the strengths and weaknesses of the M.A. programs, the study illustrated that the most satisfying aspect for students and graduates was deemed to be competent and qualified instructors of the M.A. programs. Similarly, Döner-Arday (2022), Kesli-Dollar et al. (2014), and Öztürk (2015) noted that the quality of faculty members in M.A. programs and their dedication to professional growth were widely considered strengths of M.A. programs. The findings of this study correspondingly emphasized the effectiveness of the programs in hiring qualified instructors and academic staff. In contrast, the lack of elective courses, inadequate feedback, and the repetitive nature of the courses were highlighted among program weaknesses, which were also similarly voiced by participants in Çıtak's (2021) and Döner-Arday's (2022) studies. One of the striking findings of this study was that the participants regarded the instructor's feedback as being extremely important and enlightening, while also articulating the insufficiency of feedback provided by course instructors as one of the major program weaknesses. Parallel findings were reported in Çıtak's (2021) study, where he reported "almost no critical feedback

from the academic staff related to students' progress was highlighted frequently" (p. 65). Accordingly, both graduates and students in this study likewise required more feedback and expressed their dissatisfaction in this domain. Additionally, the lack of variety in elective courses was highlighted by Ph.D. students as a concern in Bilican's (2014) study. It could be indicated that graduate programs often feature experienced faculty members who bring a wealth of knowledge and research expertise to the classroom. Hence, the development of M.A. programs in these areas is essential to provide students with better and higher-quality content. Overall, the strengths identified correspond to effective *Input* and *Process* components, while the weaknesses point to areas within the *Input* and *Process* dimensions that require attention.

Results concerning Research Question 3.2. (RQ3.2.), which delved into the students' and graduates' suggestions for the improvement and development of the graduate program they are enrolled in/graduated from, focusing on increasing course variety (particularly that of elective courses), avoiding repetitive courses, and providing adequate instructor feedback. For example, most participants suggested increasing the variety of elective courses. Similar suggestions were voiced in past studies (Bilican, 2014; Döner-Arday, 2022; Kesli-Dollar, 2014). For example, Ph.D. students in Bilican's (2014) study were primarily concerned with the lack of variety in the elective courses. Likewise, some participants in Kesli-Dollar et al.'s (2014) study requested more diverse courses on subjects such as *Introduction to Special Needs Education* and *Neurolinguistic Programming (NLP)*, even if they were content with the course variety overall. Participants in Döner-Arday's (2022) study also highlighted the need to increase the variety of courses offered per semester. Another suggestion made by our participants was to modify the course content to prevent the repetition of courses offered in undergraduate degree programs. Similar suggestions were voiced in Çitak's (2021) and Döner-Arday's (2022) studies. This finding is likely attributed to the inclusion of fundamental courses in M.A. programs, such as *Second Language Acquisition* or *Applied Linguistics*, which are also offered in undergraduate programs as a foundation for further study. Interestingly, however, while ELT B.A. graduates expressed dissatisfaction with the repetitive nature of M.A. courses, those from different departments did not voice similar issues. Participants of this study also suggested receiving instructor feedback. Specifically, the participants expressed their demands for receiving more detailed feedback on their assignments and thesis writing procedures. Viewed from a CIPP perspective, these suggestions clearly highlight the areas where the *Input* and *Process* components require systematic improvement to enhance the overall *Product*.

Conclusions, Limitations, and Suggestions for Future Research

This mixed-method descriptive study investigated the perspectives of students and graduates of M.A. programs in English Language Education (ELE) in the Turkish context. Drawing upon the findings, it can be concluded that M.A. programs are generally perceived as a fundamental first step in pursuing graduate education, as well as in fostering academic and personal growth. Participants' experiences and perspectives were also considered crucial for future academic pursuits. Overall, the study highlights the effectiveness of these programs in preparing M.A. candidates to become competent and qualified field researchers. Therefore, programs should

support students in acquiring the knowledge and skills required for their future employment and aim to enhance their satisfaction levels (Deniz-Yücer, 2020; Kırmızı, 2011; Küçüköğlu, 2015; Öztürk, 2015; Thumvichit, 2021). When interpreted within the CIPP framework, this study suggests that the programs demonstrate strengths in *Context*, such as meeting learners' motives for pursuing graduate education, and in *Product*, particularly in developing research competence and academic identity.

Still, these conclusions should be interpreted in consideration of the study's limitations. First, the participant cohort included only M.A. program students and graduates; future research could therefore incorporate additional stakeholders to ensure a more comprehensive evaluation and examine potential correspondences or discrepancies across different groups, such as instructors and administrators. Second, some participants attended their M.A. courses online due to the COVID-19 pandemic. The perspectives of students who have experienced distance education may differ from those of students in regular, in-person programs, and the potential influence of these educational types should be considered when interpreting the study's findings.

Based on the findings, several recommendations for program improvement and future research emerge. M.A. programs should aim to enhance course variety and ensure that content remains up to date, promoting intellectual enrichment while also considering students' needs, interests, and expectations. Programs should also respond to emerging trends in the field, such as the integration of Artificial Intelligence (AI), which could be offered as an elective course. Furthermore, instructor supervision should involve providing detailed and timely feedback to support graduate student growth and program quality, as inadequate feedback may limit opportunities for development. Overall, adopting a CIPP-informed lens can guide institutions in making continuous improvements and ensure that ELE M.A. programs remain responsive to the evolving needs of both academics and professionals. Accordingly, the suggestions offered here could have practical value for program developers, faculty members, and prospective M.A. candidates, contributing to the ongoing enhancement of graduate education in the field.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Acknowledgements

This article is derived from the first author's master's thesis supervised by the second author. The article has been entirely rewritten with substantial revisions extended for journal publication.

References

- Atar, C., Kır, E., & Denkci Akkaş, F. (2020). An evaluation of the English Preparatory Program at İstanbul Medeniyet University. *Language Teaching and Educational Research*, 3(1), 94–115. <https://doi.org/10.35207/later.714457>
- Bayram, I., & Canaran, Ö. (2019). Evaluation of an English preparatory program at a Turkish foundation university. *Journal of Language and Linguistic Studies*, 15(1), 48–69. <https://doi.org/10.17263/jlls.547606>

- Bilican, R. (2014). Impact of a Ph.D. ELT program on academic development of students at a Turkish state university. *ELT Research Journal*, 3(3), 111–139.
<https://dergipark.org.tr/tr/pub/eltrj/issue/5483/74459>
- Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program development*. Heinle & Heinle.
- Council of Higher Education (CoHE). (2020). *Regulations on Graduate Education*. Ankara, Turkey: CoHE.
- Creswell, J. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Pearson Education.
- Çıtak, T. (2021). *An evaluation of MA English Language Teaching programme at Çukurova University from students' perspectives* [Unpublished master's thesis]. Çukurova University.
- Deniz-Yücer, D. (2020). *An evaluation of a doctoral program in English language teaching at a Turkish university from students' perspectives by using context, input, process, product (CIPP) model* [Unpublished master's thesis]. Bahçeşehir University.
- Döner-Arday, B. (2022). *Insights from students and graduates into the Master of Arts program in English Language Teaching at Middle East Technical University* [Unpublished master's thesis]. Middle East Technical University.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2010). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Pearson.
- Foroozandeh, E., Riazi, A., & Sadighi, F. (2007). TEFL program evaluation at master's level in Iran. *Teaching English Language*, 2(2), 71–100.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2022). *How to design and evaluate research in education* (11th ed.). McGraw-Hill.
- Garton, S., & Graves, K. (2014). Identifying a research agenda for language teaching materials. *The Modern Language Journal*, 98(2), 654–657. https://doi.org/10.1057/9781137023315_1
- Kesli-Dollar, Y., Tolu, A. T., & Doyran, F. (2014). Evaluating a graduate program of English language teacher education. *Turkish Online Journal of Qualitative Inquiry*, 5(2), 1–10. <https://doi.org/10.17569/tojqi.09132>
- Kırmızı, Ö. (2011). *An evaluation of M.A. ELT programs in Turkey* [Unpublished doctoral dissertation]. Hacettepe University.
- Kuzu, E., Özkan, Y., & Bada, E. (2021). An EFL program evaluation: A case from Turkey. *The Reading Matrix: An International Online Journal*, 21(1), 94–106. <https://eric.ed.gov/?id=EJ1296261>
- Küçüköğlu, H. (2015). *An evaluation of Ph.D. in ELT programs in Turkey* [Unpublished doctoral dissertation]. Hacettepe University.
- Malagón-Amor, Á., Córcoles-Martínez, D., Martín-López, L. M., & Pérez-Solà, V. (2015). Hikikomori in Spain: A descriptive study. *International Journal of Social Psychiatry*, 61(5), 475–483. <https://doi.org/10.1177/0020764014553003>
- Mede, E. (2012). *Design and evaluation of a language preparatory program at an English medium university in an EFL setting: A case study* [Unpublished doctoral dissertation]. Yeditepe University.
- Okumuş, A., & Daloğlu, A. (2024). A programme evaluation study: The case of the English Language Teaching Programme. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 24(1), 187–216. <https://doi.org/10.17240/aibuefd.2024.-1327223>
- Öztürk, R. Ö. (2015). *An evaluation of the master's program in English language teaching at a Turkish university* [Master's thesis, Bahçeşehir University]. Turkish Council of Higher Education Theses Center. <https://acikbilim.yok.gov.tr/handle/20.500.12812/591945>
- Peacock, M. (2009). The evaluation of foreign-language-teacher education programmes. *Language Teaching Research*, 13(3), 259–278. <https://doi.org/10.1177/1362168809104698>
- Robinson, B. (2003). Evaluation, research and quality. In B. Robinson & C. Latchem (Eds.), *Teacher education through open and distance learning* (pp. 193–211). Routledge/Farmer.
- Saldaña, J. (2011). *Fundamentals of qualitative research* (2nd ed.). Oxford University Press.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Sage Publications.
- Stufflebeam, D. L. (1983). *The CIPP model for program evaluation*. Kluwer Academic Publishing.
- Thumvichit, A. (2021). Language education in emergencies: A systematic review. *Journal of Language and Education*, 7(28), 183–197. <https://doi.org/10.17323/jle.2021.12462>
- Tomlinson, B. (Ed.). (2023). *Developing materials for language teaching* (3rd ed.). Bloomsbury Publishing.
- Zeybek, G., & Tunçer, M. (2020). Feelings and problems of being a Turkish ELT graduate student: A qualitative study on graduate student perceptions. *Eurasian Journal of Teacher Education*, 1(2), 148–171. <https://dergipark.org.tr/tr/download/article-file/1151122>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the Journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (CC BY-NC-ND) (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Appendices

Appendix A

Open-ended Questions

1. What are/were you satisfied MOST about the M.A. program (e.g., content, courses, resources) you are enrolled in/graduated from? (Please explain briefly.)
2. What are/were you satisfied LEAST about the M.A. program (e.g., content, courses, resources) you are enrolled in/graduated from? (Please explain briefly.)
3. Do you have any suggestions for the improvement of M.A. programs? If yes, what are they?
4. Do you have any other comments/suggestions?

Appendix B

Interview Questions

1. What was your goal and motivation to pursue an M.A. degree?
2. What was the most influential/important factor in your decision to begin your M.A. study? Was it due to
 - 2.a. primary career choice?
 - 2.b. change of career?
 - 2.c. advanced degree required for career advancement?
 - 2.d. increased income-earning potential?
 - 2.e. personal intellectual enrichment?
 - 2.f. other?Please elaborate further.
3. In your opinion, what are/were the strengths of the M.A. program you are enrolled in/graduated from regarding
 - 3.a. the courses offered in the program?
 - 3.b. the program resources available for you?
 - 3.c. the program content?
 - 3.d. the academic staff?
4. In your opinion, what are/were the weaknesses of the M.A. program you are enrolled in/graduated from regarding
 - 4.a. the courses offered in the program?
 - 4.b. the program resources available for you?
 - 4.c. the program content?
 - 4.d. the academic staff?
5. Are/were you overall satisfied or dissatisfied with the M.A. program you are enrolled in/graduated from? Why/why not?
 - 5.1. What are/were you MOST satisfied with? Please elaborate.
 - 5.2. What are/were you LEAST satisfied with? Please elaborate.
6. What would be your suggestions for the improvement of M.A. programs in Türkiye?
7. Would you like to add any other issues than those we discussed?