An Analysis of the Studies Conducted in the Field of Curriculum Evaluation from a "Methodology" Perspective

Esra Doğan[1], Erdal Bay[2] & Bülent Döş[3]

[1] Ministry of Public Education, Kayseri, Turkey

[2] Faculty of Educational Sciences, Curriculum and Instruction, Gaziantep University, Gaziantep, Turkey

[3] Nizip Faculty of Educational Sciences, Curriculum and Instruction, Gaziantep University, Gaziantep, Turkey

Correspondence: Esra Doğan, Provincial Directorate of National Education, Kayseri, Turkey.

Abstract

This study analyzed studies done in Turkey in the context of curriculum evaluation (CE) by asking, "How is it made? The study was carried out in two stages. In the first stage, the document analysis method used 215 theses written between 1991 and 2020 on CE were analyzed according to the "thesis review form." In the second stage, depth analysis was made through semi-structured interviews with the authors (students) and the field experts (supervisors of the authors) of the theses to make the results of the first stage more understandable. Interviews were conducted with 32 participants. A maximum sampling method was used to determine the participants. The data analysis calculated percentage and frequency values for the data obtained in the first stage. In the second stage, descriptive analysis and content analysis were carried out with the MAXQDA 2020 qualitative data analysis program. The majority of theses did not employ a CE model as a consequence of the research, and the CIPP model was the most popular CE model. Many of the theses were not justified in using the CE model. Model usage increased as time passed to the present day. Many theses used quantitative models but did not explicitly state the sampling technique. Teachers were mainly used in this research as a source for data gathering, and participant numbers ranged from 10 to 50. Additionally, most studies used questionnaires and interviews as the primary data-gathering tools. All of these findings suggest that CE studies have several flaws.

Keywords: curriculum evaluation, approach, methodological

1. Introduction

Curriculum evaluation (CE) is how organizations regularly decide whether to accept, modify, or reject curricula. It is based on rigorous empirical research that assesses the program's acceptance, value, quality, effectiveness, relevance for the society in which it operates, and how well it satisfies stakeholder and policy needs. It is the process that consists of a series of planned stages in which the decision to approve, change or reject can be made (Owen, 2020; Posavac, 2016; Fitzpatrick et al., 2004; Guskey, 2000; Stufflebeam et al., 2000; Ornstein & Hunkins,1998; Melrose, 1998; JCSEE, 1994; Worthen, Borg, & White, 1993).

Evaluation can serve various purposes, including helping design policies, create programs, assess student achievement, track how education funding is used, and win over the public (Bay & Karakaya, 2006). The benefits of CE were also listed as identifying the level of goal realization, suitability of the content, suitable speed and sequence of teaching activities, offering feedback to the teacher and participants, and selecting who can participate in future programs (McCain, 2005). CE may be summed up as evaluating a program's quality and environment suitability using scientific techniques.

Curriculum Evaluation Approaches are the underlying philosophical presumptions that influence how CE is conducted (Shadish, 1998). The content of the assessment area is made up of the many approaches the evaluators have presented. Studying various assessment methodologies "is crucial for the professionalization, scientific progress, and functionalization of curriculum evaluation," according to Stufflebeam (2001), who stressed the significance of this research.

Evaluations of curricula must explain why and how the evaluation model was created. Tyler's Model, Metfessel-Michael Model, Provus' Contradiction/Difference Model, Hammond's Cube Model, CIPP Model, Alkin UCLA Model, Stake's Convenience-Probability Model, Kirkpatrick Model, Goal Independent Model, and

Stake's Responder Model are examples of standard CE models in the literature. Its model is Demirel's Analytical Model, Participatory Evaluation, Natural Evaluation, and Illuminating Model developed by Eisner Educational Criticism.

Like studies in every other discipline, CE studies begin with the question "why" and move on to the questions "how," "what," and "when." The "methodical" aspect of the investigations provides the "how" explanation. In other words, it concerns the proper method for conducting CE research. In CE investigations, specific procedures and stages are followed. These stages include needs analysis, creating a working group, defining the curriculum evaluation, the purpose of the curriculum evaluation, curriculum evaluation approaches, and models, stakeholders, budget and resources, work schedule, scope and limitations, data collection methods, deciding on data collection tools, deciding on data sources, deciding on data analysis techniques, communicating and disseminating, ensuring use, implementing, reporting, and meta-evaluation. Controls relating to ethics and standards should then be offered. However, this study did not examine the standards and ethics environment. Since CE techniques were used in this work, methodological (How?) studies were conducted to "understand" them. The scientific research and CE approach/model dimensions employed in the studies were looked at within the methodological framework.

When the research in the CE field is analyzed, a direct study on comprehending the paradigms and methodologies in the field of CE cannot be attained.The studies (Akıncı & Köse, 2021; Tok et al., 2020; Yetkiner et al., 2019; Özüdoğru, 2018; Dündar & Meriç, 2017; Koç, 2016, Kurt & Erdoğan, 2015; Gökmenoğlu, 2014) show that the descriptive analysis of CE investigations is typically carried out. A few program evaluation components were reviewed, studies were conducted with a small sample, and some were restricted to doctorate theses. CE studies were also confined to programs connected to just one course in these studies.

1.1 Purpose of the Research

This study aimed to assess the studies in a methodological context by posing the question, "How?" these studies were conducted in curriculum evaluation. In line with the primary purpose of the research, answers to the following questions were sought:

1) What is the current situation in the context of CE approaches/methods and models in studies in CE?

i) Which CE approaches/models are used?

ii) Which approaches, methods and models are used in studies conducted in the field of CE according to years and CE purpose?

iii) What are the reasons for using approaches/methods and models by those working in the field of CE?

2) What is the situation in the context of the scientific research approach used in CE studies?

i) Which research approaches are used (qualitative, quantitative, mixed)? What is the scientific research approach according to years, CE purpose, and used CE model?

ii) What is the current situation regarding the sampling methods used in CE studies? Sample size and data source, what is the current situation in the number of data sources?

iii) What is the situation regarding the limitations stated in the studies conducted in CE?

iv) What are the data collection tools used in studies in the field of CE? What is the situation in data collection tools according to CE models?

v) What are the data analysis methods used in CE studies?

vi) What is the current situation regarding the reliability and validity method techniques used in CE studies?

2. Method

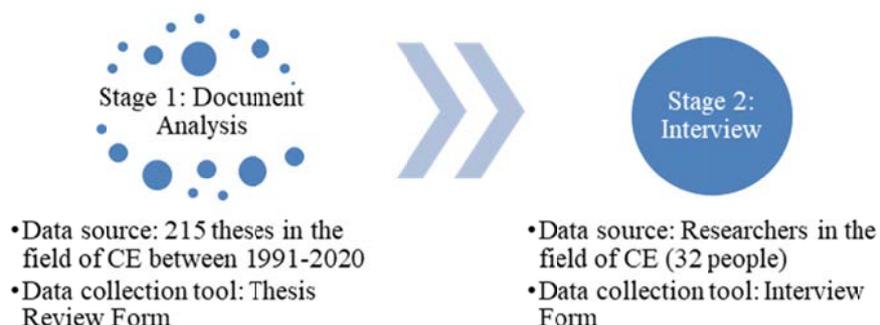The study was carried out in two stages.

Figure 1. General data collection method/data source/data collection tool model of the study

The theses on CE written between 1991 and 2020 were assessed in the first stage using the document analysis approach and the thesis review form. The second stage's goal was to clarify the first stage's findings and conduct a thorough analysis using semi-structured interviews with the theses' authors—students—and subject matter experts (advising faculty members). Data diversification was done similarly.

A. Document analysis

Meticulous and systematic examination of documents to make sense and understand the relevant subject is defined as a document analysis method (Corbin & Strauss, 2008; Wach, 2013 cited by Kıral, 2020). The following processes were carried out in document analysis.

i) Determining the criteria to be included in the documents

The focus is on postgraduate theses in the field of CE between 1991-2020. The criterion sampling method was used to determine the theses to be examined in the study. The criteria were determined that it should be made between 1991 and 2020, be open to using, the theses should be in the field of CIP, the advisors should be CIP or Educational Sciences doctorate graduates, and the download date should be 10.06.2019.

ii) Document and data collection

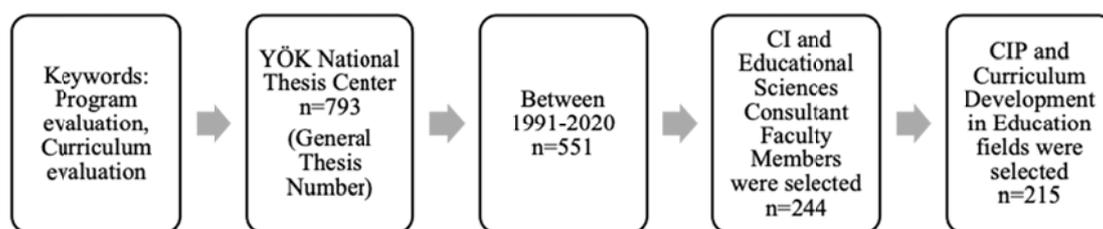The documents were tried to be reached using keywords from the YÖK thesis search page.

Figure 2. The process of reaching the theses to be examined

The study could not include many theses that were not permitted to be published or accessed on the YOK thesis screening portal. Additionally, theses that examined a program from a different nation were excluded from the study. There were 244 theses left after the adviser faculty members for CI, and educational sciences were chosen. Finally, once the thesis's field of CIP and curriculum development in education was considered during the elimination process, 215 master's and doctorate theses were left as documents that needed to be inspected.

The number of thesis studies on CE has increased during the 1991–1995 and 2011–2015 time periods; however, between 2016 and 2020, there was a decline. Seventy doctoral dissertations and 145 master's theses were employed in the investigation. Sixty-four of the theses were done at the "higher education" level, 48 at the primary school level, and 43 at the secondary school level. It was observed that n=94 of the master's theses and n=51 of the programs were related to the course/course program. It was determined that n=27 of the doctorate theses and n=43 were related to the course/course program, respectively.

It was discovered that the "English" curriculum had the highest evaluations among the course/course kinds, while

the "Ethics" curriculum received the lowest evaluations. The subject of the thesis is 228 codes for the program part, and "All" (82.9%) of these codes receive the most attention. The program's "Learning-Teaching Process" (5.3%), "Aim/Outcome" (4.8%), "Content" (3.9%), and at least "Assessment-Evaluation" (2.6%) were then the main points of attention. It can be noticed that 0.4 does not specify it.

iii) Identifying key analysis areas

The theses evaluated as part of the research were decided to be analyzed under the headings of approach/model, scientific research approach, sampling, data source, limitation, data collecting, data analysis, and validity-reliability in a "methodical" context. They were transformed into a "thesis review form."

Coding the document

Theses were examined using a "thesis review form." First, surface reading was used to check the titles of the analyzed theses during coding. The evaluation model, scientific study technique, sampling method, sample size, data source type, restriction, data collection tool type, data analysis type, and validity/reliability type were all coded in a Microsoft Excel file.

v) Verification and analysis

The research was conducted by the analytical units chosen at the start of the investigation. In this case, descriptive analysis was employed. The codes were inserted into the appropriate themes per the specified themes. The frequency and percentage values of these codes were then calculated.

B. Interview

In order to corroborate and elaborate the findings we acquired in the first part of the research in the context of data diversity and to highlight the challenges in the field, semi-structured interviews were performed with the pertinent stakeholders. So why and how? In order to get the responses, interviews were performed.

i) Preparation of the Interview Form

A semi-structured interview form was created for the research questions based on the outcomes of the pertinent literature review and the findings from the first stage. An expert's opinion was sought to guarantee the legitimacy of this interview form, and any necessary adjustments were made in response to the comments. The interview form was tested in a trial application to guarantee its authenticity. Once more, changes were made per the results, and the final interview form was received. The interview form's final version has 22 questions with five subtopics in each (Annex-1).

ii) Determining the Individuals to be Interviewed

The maximum variation sampling method—one of the purposive sampling techniques—was employed because selecting the sample for the study's interviews required a comprehensive methodology. In order to achieve a uniform distribution from both the students who authored the thesis (16) and the advisers to the thesis, the period between 1991 and 2020 was divided into five-year procedures according to the titles of the faculty members, the education level of the program in which the thesis was assessed, and whether it was a master's or doctorate thesis.

A sample of 32 people from faculty members (16) was chosen. Twenty-nine of the individuals took PD-related courses, while three did not, according to an analysis of their academic education processes. Additionally, it was found that 12 participants were exclusively applied, ten individuals were only theoretical, and 18 participants were both theoretical and practical.

iii) Application of the Interview Form

In the study's second phase, appointments were set voluntarily by phone or email, and interviews were conducted and videotaped over the phone and using the zoom application. On average, interviews took 40 to 45 minutes to complete. The interview transcripts were turned into interview texts and distributed to the participants through email or phone. The participants were requested to examine the texts, and after their review, the required portions received the necessary adjustments. During the interviews, the participants were given information regarding the subject, such as explaining what the term "purposeful context" means.

iv) Analysis of the Data in the Interview Form

Descriptive and content analysis approaches were employed for the study's interview component. Prior to analysis, the data were prepared. The Word document was dictated for each of the recorded interviews. Thesis writers (students) are identified as A1,..., and A16, respectively, while adviser faculty members (field specialists) are identified as T1...T16, respectively, using pseudonyms. The analysis was aided by the use of the MAXQDA 2020 qualitative data analysis tool. The codes were created by looking at the responses, and then the categories and

topics were determined. The advisor for the thesis and the faculty member often communicated throughout the coding process.

C. Validity-Reliability

To assure the research's objectivity, extensive interaction, depth-focused data collecting, diversification, expert evaluation, and participant confirmation were conducted. Detailed descriptions and purposeful sample techniques are advised to ensure transferability. For the research, a thorough description and purposeful sampling were employed to guarantee transferability. To ensure consistency throughout the investigation, a consistency analysis was done. One month after the initial data coding was finished, the researcher coded the quantitative and qualitative data again for this purpose. To verify confirmability, a confirmation review was completed.

3. Results

3.1 Findings Related to the First Research Question

In order to answer the first research question, we looked at which CE methodologies and models are employed in CE studies, how they are distributed among years and CE objectives, and the rationales behind why researchers choose to employ these approaches.

i) Findings on approaches and models used in CE studies

The analysis is presented in the context of the CE assessment model used in studies in CE.

Table 1. The use of the CE model in studies conducted in the field of CE

Evaluation Approach	Evaluation Model	f	%
No Model	No Model	134	62%
Purpose-Based Approach	Tyler's Model	8	
	Metfessel-Michael Model	1	
	Hammond's Cube Model	1	
	Posner	1	
	Provus' Contradiction/Difference Model	1	
	Total	12	6%
CIPP Model	CIPP Model	29	
	Kirkpatrick Model	2	
	Alkin's UCLA Model	1	
	Total	32	15%
Expert-Focused Approach	Eisner Educational Criticism Model	6	3%
Participant-Focused Approach	Staking Availability-Probability Model	2	
	Participatory Evaluation	2	
	Bellon and Handler Model	2	
	Stake (Responsive) Model	1	
	Illuminating Model	1	
	Demirel's Analytical Model	1	
	Stake (Standard-Based) Model	1	
	Total	10	5%
Others	Undefined, Investigative Model, Program Elements, Model Suggestion, Total, Product and Process, Bloom, Delphi, Karma, Lawler and King's Adult Education Model, Pombo and Moreira's categories	20	9%
	Total	215	100%

Table 1 analysis revealed that, out of 215 theses, 134 (62%) did not employ the assessment model. It was observed

that the CIPP model (15%) and "management-based techniques" (14%) were utilized most frequently in theses. It has been observed that theses frequently use methodologies, strategies, and expressions for CE that are unrelated to CE approaches and models, including delphi, bloom, total, product, and process.

ii) Findings regarding the approaches/methods and models used in studies conducted in the field of CE by years and type of CE purpose

The analysis in the context of CE evaluation approaches/methods and models used periodically by years in studies in the field of CE is presented in Table 2.

Table 2. CE approach/methods and models by years

		1991-1995	1996-2000	2001-2005	2006-2010	2011-2015	2016-2020	Toplam
Management Based Approach	UCLA	0	0	0	0	1	0	1
	CIPP	0	0	1	3	9	16	29
	Kirkpatrick	0	0	0	1	0	1	2
Purpose-Based Approach	Tyler	0	0	0	1	5	2	8
	Hammond's Cube	0	0	0	0	0	1	1
	Metfessel Michael	0	0	0	0	0	1	1
	Posner	0	0	0	0	1	0	1
	Proof	0	0	0	0	1	0	1
Expert-Focused Approach	Eisner Educational Criticism	0	0	1	1	2	2	6
Participant-Focused Approach	Stake's (Convenience-Probability)	0	0	0	0	2	0	2
	Stake (Standard-Based)	0	0	0	1	0	0	1
	Stake (Responsive)	0	0	0	1	0	0	1
	Participatory Model	0	0	0	1	1	0	2
	Illuminating Model	0	0	0	0	0	1	1
	Bellon & Handler	0	0	0	0	0	2	2
Other	-	0	2	4	2	8	7	23
No model	-	1	3	6	41	61	22	134

The most significant intersection in Table 2 was between the 2011–2015 time period and the statement "no model." This intersection can be attributed to the fact that models were not used in 134 theses. Additionally, it was discovered that from 1991 to 1995, no models were utilized and that between 2016 and 2020, when 34 researches were conducted, most models were employed. Additionally, it is significant that studies typically use the CIPP model, but alternative models are occasionally used.

Table 3 analyzes the CE evaluation approaches, methods, and models utilized in CE research according to the goal type.

Table 3. CE approaches/methods and models according to purpose type in studies conducted in the field of CE

		Instrumental	Comparison	Decision Making	Idealization	Essentialist	Total
Management Based Approach	UCLA	0	1	0	1	1	3
	CIPP	11	2	4	15	27	59
	Kirkpatrick	1	0	0	0	2	3
Purpose-Based Approach	Tyler	5	0	0	1	5	11
	Hammond's Cube	1	0	0	0	0	1
	Metfessel Michael	0	0	0	0	1	1
	Posner	0	0	0	1	1	2
	proof	0	0	0	0	1	1
Expert-Focused Approach	Eisner Educational Criticism	2	1	0	3	6	12
Participant-Focused Approach	Stake's (Convenience-Probability)	2	0	0	0	2	4
	Stake (Standard-Based)	0	0	0	0	1	1
	Stake (Responsive)	1	0	0	1	1	3
	Participatory Model	1	0	0	1	2	4
	Illuminating Model	1	1	0	1	1	4
	Bellon & Handler	1	0	0	1	2	4
Other		11	3	3	9	20	46
No Model		17	6	12	25	121	181

The essentialist inquiry type and the no model statement have shown the most significant amount of intersection in Table 3. The UCLA approach employed comparison, idealization, and essentialist questions kinds. There was no use of a decision-making question type. The CIPP model used the most essentialist question type but hardly ever used the question type with decision-making utility. Instrumental and essentialist question types were most frequently utilized in the Tyler model.

iii) Findings regarding the reasons for using approaches and methods/models of CE students

The CE approach used by those working in the field of CE and the reasons for using it is presented in Table 4.

Table 4. The CE approach used by researchers who conducted CE studies and their reasons for using it

Which Approach	Rationale Code
Purpose-based (n:13/43%)	It is important to reach the goals (n=1)
Based on management (n:6/19%)	Clear planning (n=1)
Participant-oriented (n:5/16%)	Applicable (n=1)
Expert focused (n:4/13%)	-
Many different approaches with the team (n:2/6%)	Gaining importance with constructivism (n=1)
Consumer focused (n:1/3%)	No evaluation alone (n=1)
Total n:31/100%	Being a professional process (n=1)

According to the interviews with researchers working in the field of CE, when Table 4 was evaluated, 32 researchers provided 31 codes. The majority of techniques are "purpose-based," followed by "management-based," "participant-oriented," "expert-oriented," and "consumer-oriented."

However, it was discovered from examining the table that many researchers cannot defend their methods. There were a total of 8 codes mentioned by those who gave their justifications. It can be seen from the table that the researchers take these techniques into account for their studies since, for instance, utilizing a goal-based approach is crucial to achieving the aims.

In this regard, "A10," one of the researchers, claims that he employed the "purpose-based" methodology in the following manner:

> *"I particularly tried to do goal-oriented CE because it is important to achieve goals, especially in the education program I studied, which is aimed at gaining communicative competence and at students becoming proficient in all of their reading, writing, speaking, and listening skills" (A10 Interview Transcript, Position 42).*

Upon examining the A10 opinion, it is apparent that the author uses a purpose-driven approach in his/her thesis. However, he fails to state why he holds this opinion.

When the justifications listed in the table are taken into consideration, it becomes clear that "it is crucial to achieving the goals" for the goal-based approach, "it gains importance with constructivism" for the participant-oriented approach, "it is a professional process" for the expert-oriented approach, and "to expose how the program is processed" for many different methods in collaboration with the team. Gave their justifications.

The findings related to the CE model used by the researchers who conducted CE studies and the reasons for using this model are presented in Table 5.

Table 5. The CE model used by researchers working on CE and the reasons for using it

Which model	Justification Code
CIPP Model	Being a flexible model (n=1)
(n:13/23%)	Being the most suitable model for the participant (n=1)
No models	Being functional (n=1)
(n:10/19%)	Clear planning (n=1)
Program comparison	Being very comprehensive (n=6)
(n:5/%9)	Models are not valid today (n=1)
Kirkpatrick Model	Models are special examples of Input-process-output structure (n=1)
(n:4/7%)	Looking at the descriptive statistics of achievements and activities (n=1)
Provus' Model of Differences	Using models is difficult with graduate students (n=1)
(n:3/6%)	-
Tyler's Model	Since it is the most suitable model for in-service training (n=1)
(n:3/6%)	It was a program we developed ourselves (n=1)
Stake's Eligibility-Probability Model	Being a breeding-based model (n=1)

According to the data gathered from the interviews with academics working in the field of CE, 32 researchers provided 54 codes when Table 5 was analyzed. The "CIPP" model was most frequently used, and it was followed by "No model," "Kirkpatrick," "Provus' Differences Model," "Tyler's Model," "Stake's Suitability-Probability Model," "Alkin's Model," "Hammond's Cube Model," "Demirel's Analytical Model," "Eisner Educational Criticism Model," "Illuminating Model," "Target-free Model,"

However, a closer examination of the table reveals that many researchers cannot justify why they use these models. Those who presented their justifications also stated a total of 22 codes.

The table shows that the researchers' motivation for considering these models in their studies, such as the CIPP model, is "to be very comprehensive".

In this regard, one of the researchers states that he used the "CIPP" model in his study "A16" as follows:

> *"I used the CIPP model. I used it because it offers the opportunity to evaluate each step of the program separately and is quite comprehensive." (A16 interview form, Position 42).*

When the A16 opinion was analyzed, it was discovered that the author employed the CIPP model in his thesis and articulated it in this manner, but he did not provide a reasonable explanation for why.

The most frequent justifications for the CIPP model when the justifications listed in the table are examined are "to be very comprehensive," "the models are not valid today" for no model, "a model suitable for study" for the Kirkpatrick model, "guidance" by Hammond for the cube model, Demirel's argument, and "a model suitable for study" for the Kirkpatrick model. For Eisner's analytical model and "being systematic" for his educational criticism model, they gave the justifications for "being in Turkey" and "being systematic," respectively.

3.2 Findings Related to the Second Research Question

In the second research question, the analysis of the studies in the field of CE was presented in the context of the scientific research approaches used.

i) Findings on scientific research approaches used in CE studies

Findings related to scientific research approaches used in CE studies are presented in Table 6.

Table 6. Scientific research approaches used in CE studies

Scientific Research Approach	f	%
Quantitative	126	58
Qualitative	42	20
Mixed	47	22
Total	215	100

The quantitative approach was used in 58% (126) of the 215 theses examined in Table 6; 22% (47) of them used the mixed approach; It is seen that 20% (42) used the qualitative approach.

Findings of scientific research approach according to year periods in studies in the field of CE

The findings of scientific research approaches according to year periods in studies conducted in CE are presented in Table 7.

Table 7. Values of scientific research approach according to year periods in studies conducted in the field of CE

Year/Scientific Research Approach	Qualitative	Quantitative	Mixed
1991-1995	0	1	0
1996-2000	0	6	0
2001-2005	2	9	1
2006-2010	7	37	6
2011-2015	20	49	22
2016-2020	14	23	18
Total	43	125	47

The quantitative method for 2011 to 2015 shows the most significant overlap, according to Table 7. Most thesis research using quantitative, qualitative, and mixed methods was completed between 2011 and 2015.

Findings on scientific approaches used according to the aims of studies in the field of CE

The findings of CE models according to scientific research approaches in studies conducted in CE are presented in Table 8.

Table 8. Distribution of CE models according to scientific research approaches in studies conducted in the field of CE

		Qualitative	Quantitative	Mixed	Total
Management Based Approach	UCLA	1	0	0	1
	CIPP	8	10	11	29
	Kirkpatrick	0	0	2	2
Purpose-Based Approach	Tyler	0	5	3	8
	Hammond's Cube	0	0	1	1
	Metfessel Michael	1	0	0	1
	Posner	0	1	0	1
	Provus	0	1	0	1
Expert-Focused Approach	Eisner	3	2	1	6
Participant-Focused Approach	Stake's (Convenience-Probability)	1	0	1	2
	Stake (Standard-Based)	0	0	1	1
	Stake (Responsive)	1	0	0	1
	Participatory Model	0	0	2	2
	Illuminating Model	0	0	1	1
	Bellon & Handler	0	0	2	2
Other		7	9	7	23
No Model		21	98	15	134

The qualitative approach and the phrase "no model" have been found to intersect most frequently when Table 8 is evaluated. The UCLA model-using study is an empirical investigation. The CIPP model was mainly used in mixed research. Quantitative research has typically employed the Eisner model. Most importantly, qualitative research has been used to apply the Tyler model.

ii) Findings on sampling methods used in studies in the field of CE

The findings regarding the sampling methods used in the studies conducted in the field of CE are presented in Table 9.

Table 9. Distribution according to the sampling method used in studies conducted in the field of CE

Sampling Method	f	%
Unspecified	53	18
Simple Random Sampling	50	17
Purposeful Sampling-Criteria Sampling	39	13
Purposeful sampling-Maximum diversity sampling	31	10
Easily Accessible Situation Sampling	27	9
Directly to the universe	22	8
Cluster sampling	22	8
Stratification	20	7
Purposeful Sampling	8	3
Purposeful Sampling - Typical case sampling	7	2
Purposeful Sampling-Snowball Sampling	5	2
Purposeful Sampling - Simulated Sampling	3	1
Convenient Sampling	2	1
Other	2	1
Total	291	100.0

Table 9 shows that there are 291 codes linked to the sampling method in the theses, with 18% of these codes being "unspecified." It can be shown that "basic random sampling" (17%) is the most chosen sample technique. It may be shown that at least other sampling techniques are employed.

Findings regarding the sample size used in studies in the field of CE

Findings regarding the sample size used in studies conducted in the field of CE are presented in Table 10.

Table 10. Distribution by sample size used in studies conducted in the field of CE

Sample Class	n	%
1-10	145	25
11-50	169	29
51-100	66	11
101-300	97	16
301-500	52	9
501-1000	44	8
1001 and above	11	2
Total	584	100.0

The sample size between "11-50" (29%) is the most frequently used among the 584 codes relating to the sample size in the theses, according to Table 10. As observed, at least a sample size of "1001 and above" (2%) is being used.

Findings related to the type of data source used in studies in the field of CE

Findings related to the type of data source used in studies conducted in the field of CE are presented in Table 11.

Table 11. Distribution by type of data source used in studies conducted in the field of CE

Data Source	n	% n	% Thesis
Teacher	136	32	63
Student (university)	49	11	23
Student (primary education)	46	10	21
Teaching staff	48	11	22
Executive	20	5	9
Syllabus	32	8	15
Student (high school)	9	2	4
Lecturer	6	1	3
Graduate	7	2	3
VK_other_person	43	10	20
VK_other_dump	33	8	15
Total	429	100	

Table 11 shows that there are 429 codes connected to the data source in the theses, with "teacher" accounting for 32% of these codes. The group with the fewest students is the "lecturer" (1%).

Findings on the number of data source types used in studies in the field of CE

The findings regarding the number of data source types used in studies conducted in the field of CE are presented in Table 12.

Table 12. Distribution by the number of data source types used in studies conducted in the field of CE

Data Source	n	%
1	103	48
2	55	26
3	28	13
4	18	8
5	7	3
6	3	1
7	1	1
Total	215	100

In 48% (103) of the 215 theses examined in Table 12, one; 26% (55) had two; 13% (28) had three; 8% (18) had four; 3% (7) five; six (3) in 1%; 1% of (1) seven; data source type is used.

iii) Findings on the limitations of studies in the field of CE

Findings related to the limitations of studies conducted in the field of CE are presented in Table 13.

Table 13. Distribution according to limitations in studies conducted in the field of CE

Limitations	f	%
No Model	134	40
No data variation	87	26
Only Educator Opinion	75	23
Only Learning Vision	16	5
Single model	8	2
Small sample size	8	2
Single unit	2	1
Other	2	1
Total	332	100.0

Table 13 reveals that there are 332 codes connected to the theses' constraints, with "no model" (40%) being the most restrictive of these categories. It can be seen that there are, at the very least, "single unit" and "other" constraints (only the administrator's opinion and only the parent's opinion, respectively) (1%).

iv) Findings on the type of data collection tool and data collection tool used in studies in the field of CE

The type of data collection tool used in CE studies and the findings regarding the data collection tool is presented in Table 14.

Table 14. Distribution by type of data collection tool and data collection tool used in studies conducted in the field of CE

Data Collection Tool Type	Data collection tool	f	%
Meeting	Interview form	201	31
	Focus Group Interview	23	3
Questionnaire	Questionnaire	184	28
	Needs Analysis Survey	6	1
	Open Ended Survey	3	1
Observation	Observation Form	53	8
Documents	Document	49	8
	Daily	4	1
Scales	Attitude Scale	4	1
	Scale	44	6
Tests	Achievement test	17	3
	Tracking Test	6	1
	Access Test	4	1
Forms	Student Evaluation Form	6	1
	Program Evaluation Form	4	1
Other	Other Data Collection Tools	34	5
	Total	642	100

When Table 14 is examined, it is seen that there are 642 codes related to data collection tools in theses, and among these codes, "interview" (31%) is used as the data collection tool type. It is seen that forms are used at least.

Findings of CE models according to the data collection tool in the studies in the field of CE

The findings of CE models according to the data collection tool included in the studies conducted in the field of CE are presented in Table 15.

Table 15. Distribution of CE models according to the data collection tool in CE studies

		Interview form	Questionnaire	Observation Form	Document	Scale	Focus Group Interview	Achievement test	Other
Management Based Approach	UCLA	1	0	0	1	0	0	0	0
	CIPP	21	11	7	11	6	8	2	5
	Kirkpatrick	2	1	1	0	0	0	0	1
Purpose-Based Approach	Tyler	5	4	1	0	0	0	6	4
	Hammond's Cube	1	0	1	0	1	0	0	1
	Metfessel Michael	1	1	1	1	0	1	0	0
	Posner	0	1	0	0	0	0	0	0
	Provus	1	0	0	0	0	0	0	1
Expert-Focused Approach	Eisner	5	2	1	2	0	2	1	2
Participant-Focused Approach	Stake's (Convenience-Probability)	2	0	2	1	0	0	1	1
	Stake (Standard-Based)	1	1	0	0	0	0	0	0
	Stake (Responsive)	0	1	0	1	0	1	0	1
	Participatory Model	2	2	1	1	0	1	0	0
	Illuminating Model	1	0	1	1	1	0	1	1
	Bellon & Handler	2	2	1	2	1	0	0	0
Other		15	12	6	11	5	2	3	3
No Model		51	79	14	16	23	3	2	18

Table 15 analysis revealed that the questionnaire and no model had the most outstanding amount of intersection. The study, which applied the UCLA model, used interview forms and materials. The interview form was employed more frequently than the achievement test in the research utilizing the CIPP paradigm. The Eisner model employed the interview form more frequently than the achievement test and observation form. The achievement test and observation form were used sparingly in research based on the Tyler model.

v) Findings related to data analysis methods in studies conducted in the field of CE

Findings related to data analysis methods in studies conducted in the field of CE are presented in Table 16.

Table 16. Distribution according to data analysis methods in studies conducted in the field of CE

Method	Data Analysis Method	f	%
Descriptive Statistics	Descriptive It.	256	24
	Content Analysis	250	23
	Descriptive Analysis	128	12
Inferential Statistics	T-test	123	11
	ANOVA	108	10
	Mann Whitney U	60	6
	Kruskal Wallis	49	5
	Chi-square	34	3
	Correlation	15	1
	Wilcoxon Signed Rank Test	10	1
	ANCOVA	9	1
	Tetrachoric Correlation	6	1
	Other Inferential Statistics Methods	26	2
Total	Total	1074	100.0

When Table 16 is studied, it can be observed that there are 1074 codes relating to data analysis methods in the theses, with the majority of these codes (59%) using the descriptive statistics data analysis method. Other research has demonstrated adopting the "inferential statistics" data analysis method (41%).

vi) Findings on the number of validity types in studies conducted in the field of CE

The findings regarding the number of validity types included in studies conducted in the field of CE are presented in Table 17.

Table 17. Distribution according to the number of validity types in studies conducted in the field of CE

Validity	f	%
-	14	6
1	38	18
2	53	25
3	50	23
4	19	9
5	17	8
6	11	5
7	9	4
8	4	2
Total	215	100

The validity type was not utilized in 6% (14) of the 215 theses that were looked at in Table 17. Instead, 18% (38) had one, 25% (53) had two, 23% (50) had three, 9% (19) had four, 8% (17) had five, below 5% (11), and 4% (9) ate. It is also noticeable that 2% (4) use eight varieties of validity.

Findings Related to Types of Validity in Studies in the Field of CE

The findings regarding the types of validity in the studies conducted in the field of CE are presented in Table 18.

Table 18. Distribution by types of validity in studies conducted in the field of CE

Validity	f	%
Expert Opinion	197	32
Pilot Application	148	23
Factor Analysis	50	8
Direct Quote	42	7
Variation	40	7
Detailed Description	38	6
Item Analysis	30	5
Participant Confirmation	23	4
Long Term Interaction	19	3
Triangulation	13	2
Depth-Oriented Data Collection	10	2
Other	6	1
Total	616	100

There are 616 codes connected to the type of validity in the theses, and "expert opinion" (32%) is the most frequently used. This information is shown in Table 18. We can see that "other" validity kinds (1%) are utilized the least frequently.

The number of reliability types included in studies conducted in the field of CE, related findings

The findings related to the number of reliability types included in studies conducted in the field of CE are presented in Table 19.

Table 19. Distribution according to the number of reliability types in studies conducted in the field of CE

Reliability	f	%
-	63	29
1	104	48
2	34	16
3	13	6
4	1	1
Total	215	100

Reliability type was not used in 29% (63) of the 215 theses examined in Table 19; one in 48% (104); 16% (34) had two; 6% (13) had three; It is seen that 0.5% (1) use four reliability types.

Findings related to reliability types in studies conducted in the field of CE

Findings related to reliability in studies conducted in the field of CE are presented in Table 20.

Table 20. Distribution by types of reliability in studies conducted in the field of CE

Reliability	f	%
Cronbach Alpha	113	53
Intercoder Consistency	65	30
Reliability_other	15	7
KR-20	13	6
Correlation	9	4
Total	215	100

Table 20 shows that Cronbach's Alpha (53%) reliability type is used the most, and correlation (4%) reliability type is used the least.

4. Discussion

CE studies were analyzed "methodologically," which corresponds to the answer to the question "How?" We were seeking answers to how CE research approaches and models were applied.

Based on the document analysis, more than half of the theses examined do not use CE models, while CIPP models are primarily used. Özüdoğru (2018), Kurt and Erdoğan (2015), and Taş and Duman (2020) concluded from their studies that most theses did not use models, and the CIPP model was prevalent in those that did. Akıncı and Köse (2021) examined the studies using the model, and it was determined that the CIPP model was most preferred. Considering that researchers mostly use the CIPP model, it is reasonable to assume that other researchers have adopted it. It was given in the findings that the expressions that are not related to the CE model in the examined theses are also used as the CE model type. This shows that some researchers lack cognitive awareness about CE models.

From a year perspective, 1991-1995 had no models, whereas 2016-2020 had the most models. CE researchers gradually become more cognizant of CE models with each passing day. According to the question type, While the most essentialist question type was used in the CIPP model, the question type with decision-making value was used very little. In the Tyler model, instrumental and essentialist question types were used the most. However, CIPP is a management-based model, and questions with decision value should be used more. The Tyler model is goal-based, and according to this approach, the outputs are more important; it can be said that question types with instrumental value can be used in this model.

The theses under review also use phrases unrelated to the CE model.

The interviews' findings show that participants primarily employ the "purpose-based" CE method in their research. The goal-based CE strategy is said to be employed for many reasons, including that "it is important to attain the goals, the planning is explicit, and it is applicable."

They employ "many ways based on management, participant-oriented, expert-oriented, together with the team," despite poor rates. It can be noted that participants primarily employ the CIPP model in their research.

The CIPP model is employed for various reasons, including that it is "comprehensive, it is a flexible model, it is the most acceptable model for the participant, it is functional, and the planning is transparent." No model, however,

program comparison, Kirkpatrick, Provus' difference, Tyler's, Stake's fitness-probability, Alkin's UCLA, Hammond's Cube, Demirel's Analytical, or Eisner Educational Criticism model, depends on context, illuminating evaluation, features of many models, Bellon and Handler's, Target-independent, or hybrid model.

Most respondents who chose the model said they did so because it was "extremely complete." They said the participants' response, "There is no model," was given because "models are specific examples of the input-process-output structure."

Akıncı and Köse (2021) examined the studies using the model related to curriculum evaluation and found that the CIPP model was preferred the most, and the reason for the "multidimensionality" was presented as the reason for choosing CIPP. Christie and Lemire (2019) also emphasized that models are theories that draw attention to the context-mechanism-outcome configuration.

It is seen that the reasons why the participants use the CE approaches and models are not clear enough, and some of the reasons do not fully meet the relevant CE approach and model.

The study observed that the code "depends on the context" was obtained about which model the participants used. Nouraey et al. (2020) concluded that selecting an appropriate evaluation model will depend on several criteria, such as the evaluation's context, purpose, and expected outcome.

It can be seen that the participants agreed that every aspect of CE is crucial. This indicates that the participants believe no component is more significant than the others.

Document analysis reveals that more than half of the theses studied as documents in the "methodological" context take a quantitative approach; this correlates to the response to the "how" issue in CE studies. It was found that the most quantitative research methodology was applied in the studies of Koç (2016) and Kurt and Erdoğan (2015). Kurt and Erdoğan (2015) interpreted this to mean that the quantitative approach's prominence is because it yields findings more quickly than other approaches.

Again, Azzam (2011) discovered that the evaluators chose a more quantitative methodology in his study, which studied the impact of specific demographic characteristics of the evaluators on their choice of methodology. This result supports the quantitative approach to taking the lead.

Most of these studies lack the characteristics of formal CE studies; in real CE studies, philosophy, ideology, question type, and more qualitative studies should predominate. When CE is generally mentioned in studies and success tests, there is logic to obtaining numerical data by using CE, but it is seen that the quantitative approach is more common.

In the study of Yetkiner et al. (2019), the study of Tok et al. (2020), the study of Akıncı and Köse (2021), and the study of Özüdoğru (2018), the number of mixed studies were found to be higher. Some researchers also state that paradigms tend to coexist and are used when appropriate; for example, they argue that while some research questions are suitable to be analyzed by quantitative methods, some questions are suitable to be analyzed by qualitative methods (Erickson, 1986; Lakatos, 1978). This may explain the greater use of mixed studies in these studies.

It can be observed from the year periods that the years 2011 to 2015 are when the quantitative technique is most frequently applied. The research that used the most qualitative method did not use the model according to the PD models. A quantitative approach can be applied in graduate studies between 2011 and 2015.

It is common to find that the sampling technique employed in CE studies is "basic random sampling," which is not stated in CE studies to a great extent. Similar findings were observed in the investigations of Kurt and Erdoğan (2015) and Özüdoğru (2018), where the majority of the studies did not specify the sampling method and the random (simple random sampling) method predominated. The bulk of the studies that Akıncı and Köse (2021) looked at failed to mention the sampling technique. It was observed that the population received the most significant direct application in the studies listed.

The study's sample size range was found to be, at most, 11 to 50 when it was investigated. On the other hand, Kurt and Erdoğan (2015) discovered that it was more prevalent in the 301–1000 range, and Özüdoğru (2018) observed that it was more prevalent in the 31–100 range. These findings conflict with what we found.

The most prevalent data source type is "teacher," and more than half of the research employs multiple data sources. The teacher is the prominent data source in the study. The teacher is the most prominent data source in Kurt and Erdoğan's (2015) studies. In the study of Koç (2015), the teacher came to the fore. The study by Özüdoğru (2018) determined that teachers/trainers are the majority.

Since teachers administer programs and are a significant stakeholder group in CE studies, it can be hypothesized

that this is why more teachers are chosen as data sources. In the studies of Yetkiner et al. (2019), it was seen that academics were more prominent. In the studies of Taş and Duman (2020), it was determined that most documents were used as data sources.

There is no model in more than half of the investigations, which is a restriction. The most popular methods for gathering data are questionnaires and interviews. In studies that do not employ models, it is evident from the models that surveys are mainly used.

The two techniques most frequently utilized in data analysis are descriptive statistics and content analysis. Descriptive statistics was the study's most popular data analysis technique, followed by content analysis. Kurt and Erdoğan (2015) found that descriptive statistics are at the forefront of the quantitative method and content analysis is in the foreground of the qualitative method. In the study of Özüdoğru (2018), it was seen that descriptive statistics were also in the foreground.

It can be shown that more than half of the research used multiple types of validity, with the expert opinion being the most prevalent. In more than half of the studies, reliability was not specified, or only (1) reliability type was stated. It is seen that Cronbach Alpha reliability is the most used reliability type. In the study, it was determined that expert opinion came to the fore in validity methods. Özüdoğru (2018) found that expert opinion was used the most as a validity method in the theses he examined. As for reliability, it was seen that Cronbach's Alpha method was mainly used.

In light of this, it can be argued that the evaluation approach and model used in research undertaken in CE have flaws and shortcomings. In conclusion, it can be claimed that the methodological framework of the studies done in CE has issues and deficiencies.

References

Akıncı, M., & Köse, E. (2021). 2010-2019 yılları arasında türkiye'de program değerlendirme konusunda yapılan araştırmaların eğilimleri. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi, 50*(1), 77-120.

Azzam, T. (2011). Evaluator characteristics and methodological choice. *American Journal of Evaluation, 32*(3), 376-391. https://doi.org/10.1177/1098214011399416

Bay, E., & Karakaya, Ş. (2006). Program değerlendirme sürecinde karşılaşılan sorunlar. *Kazım Karabekir Eğitim Fakültesi Dergisi*, 13.

Christie, C. A., & Lemire, S. T. (2019). Why evaluation theory should be used to inform evaluation policy. *American Journal of Evaluation, 40*(4), 490-508. https://doi.org/10.1177/1098214018824045

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Thousand Oaks: Sage. https://doi.org/10.4135/9781452230153

Dündar, E., & Merç, A. (2017). A critical review of research on curriculum development and evaluation in ELT. *European Journal of Foreign Language Teaching, 2*(1), 136-168.

Erickson, F. (1986). Qualitative methods in research on teaching. In M. Wittrockk (Ed.), *Handbook of Research on Teaching* (3rd ed., pp. 119-161). New York: MacMillan.

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). *Program evaluation: Alternative approaches and practical guidelines*. Boston: Allyn and Bacon. Pearson Education.

Gökmenoğlu, T. (2014). Broad perspective: curriculum evaluation studies in Turkey in terms of models and approaches. *International Journal of Curriculum and Instructional Studies, 4*(7), 54-70.

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press, Inc.

Kıral, B. (2020). Nitel bir veri analizi yöntemi olarak doküman analizi. *Siirt Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 8*(15), 170-189.

Koç, E. S. (2016). Türkiye'de ilköğretim programlarının değerlendirilmesine yönelik yapılan lisansüstü tezlerin incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 16*(1), 198-216. https://doi.org/10.17240/aibuefd.2016.16.1-5000182918

Kurt, A. & Erdoğan, M. (2015). Program değerlendirme araştırmalarının içerik analizi ve eğilimleri; 2004-2013 yılları arası. *Eğitim ve Bilim Dergisi, 40*(178), 199-224. https://doi.org/10.15390/EB.2015.4167

Lakatos, I. (1978). *The methodology of scientific research programmes (Volume 1): philosophical papers*. Cambridge University Press. https://doi.org/10.1017/CBO9780511621123

McCain, D. V. (2005*). Evaluation basics*. Alexandria: American Society for Training & Development.

Melrose, M. (1998) Exploring paradigms of curriculum evaluation and concepts of quality. *Quality in Higher Education, 4*(1). 37-43. https://doi.org/10.1080/1353832980040105

Nouraey, P., Al-Badi, A., Riasati, M. J., & Maata, R. L. (2020). Educational program and curriculum evaluation models: a mini systematic review of the recent trends. *Universal J Educ Res, 8*(9), 4048-4055. https://doi.org/10.13189/ujer.2020.080930

Ornstein, A. C., & Hunkins, F. P. (1998). *Curriculum foundations, principles and issues* (3th ed.). Allyn and Bacon. Boston.

Owen, J. M. (2020). *Program evaluation, forms and approaches* (3rd ed.). New York, Routledge. https://doi.org/10.4324/9781003116875

Özüdoğru, F. (2018). Analysis of curriculum evaluation studies conducted in foreign language education: 2005-2016. *Journal of Language and Linguistic Studies, 14*(2), 113-134. Retrieved from https://dergipark.org.tr/tr/download/article-file/650482

Posavac, E. J. (2016). *Program evaluation, methods and case studies* (8th ed.). New York, Routledge. https://doi.org/10.4324/9781315664972

Stufflebeam, D. L., Madaus, G. F., & Kellaghan, T. (2000). *Evaluation models* (Rev. ed.). Boston: Kluwer.

Taş, İ. D., & Duman, S. N. (2021). A systematic review of postgraduate theses on curriculum evaluation. *Uluslararası Eğitim Programları ve Öğretim Çalışmaları Dergisi, 11*(1), 43-64. https://doi.org/10.31704/ijocis.2021.003

Tok, Ş., Aslan, B., Çimenli, R., & Aksakal, Ş. N. (2020, Aralık). *Program değerlendirme çalışmalarına yönelik yapılan çalışmaların doküman inceleme çalışması (2001-2020)*. 3. Uluslararası Eğitim Araştırmaları Kongresi, IDU, İzmir. Retrieved from http://icer.idu.edu.tr/wp-content/uploads/2020/12/Fullpaper2812.pdf

Wach, E. (2013). *Learning about qualitative document analysis*. Retrieved from https://opendocs.ids.ac.uk/opendocs/bitstream/handle/20.500.12413/2989/PP%20InBrief%2013%20QDA%20FINAL2.pdf?sequence=4

Worthen, B. R., Borg, W. R., & White, K. (1993). *Measurement and evaluation in the schools* (1st ed.). Longman.

Yetkiner, A., Acar, Erdol, T., & Ünlü, Ş. (2019). Content analysis of phd dissertations on curriculum evaluation (1996-2017). *Erzincan University Journal of Education Faculty, 21*(1), 247-269. https://doi.org/10.17556/erziefd.443298

Copyrights