

DEVELOPING A PARTICIPATORY MODEL FOR TEACHER EVALUATION
AND PROFESSIONAL DEVELOPMENT: A CASE STUDY

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF SOCIAL SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

EMEL GÜNEŞ SAVUL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE DOCTOR OF PHILOSOPHY
IN
THE DEPARTMENT OF EDUCATIONAL SCIENCES, CURRICULUM AND
INSTRUCTION

OCTOBER 2022

Approval of the thesis:

**DEVELOPING A PARTICIPATORY MODEL FOR TEACHER
EVALUATION AND PROFESSIONAL DEVELOPMENT: A CASE STUDY**

submitted by **EMEL GÜNEŞ SAVUL** in partial fulfillment of the requirements for
the degree of **Doctor of Philosophy of Educational Sciences, Curriculum and
Instruction, the Graduate School of Social Sciences of Middle East Technical
University** by,

Prof. Dr. Sadettin KİRAZCI
Dean
Graduate School of Social Sciences

Prof. Dr. Hanife AKAR
Head of Department
Department of Educational Sciences

Assoc. Prof. Dr. Yeşim ÇAPA-AYDIN
Supervisor
Department of Educational Sciences

Examining Committee Members:

Prof. Dr. Fatma HAZIR BIKMAZ (Head of the Examining Committee)
Ankara University
Department of Educational Sciences

Assoc. Prof. Dr. Yeşim ÇAPA-AYDIN (Supervisor)
Middle East Technical University
Department of Educational Sciences

Prof. Dr. Cennet ENGİN
Middle East Technical University
Department of Educational Sciences

Prof. Dr. Soner YILDIRIM
Middle East Technical University
Computer Education & Instructional Technology

Assoc. Prof. Dr. Berna ASLAN
Ankara University
Department of Educational Sciences



I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name: Emel Güneş Savul

Signature:

ABSTRACT

DEVELOPING A PARTICIPATORY MODEL FOR TEACHER EVALUATION AND PROFESSIONAL DEVELOPMENT: A CASE STUDY

GÜNEŞ SAVUL, Emel

PhD. Department of Educational Sciences

Supervisor: Assoc. Prof. Dr. Yeşim ÇAPA-AYDIN

October 2022, 395 pages

This case study aimed to develop a participatory teacher evaluation model to promote teachers' professional development. A private primary school was determined as the case, and the study was conducted in the academic year of 2018-2020. Classroom teachers, principals, and experts working in this school constituted the participants.

The first part of the research aimed to provide an in-depth analysis of perceptions of participants on current teacher evaluation and professional development practices. In the second part of the study, the researcher developed an initial model in line with the opinions obtained from the first part. Initial model was introduced to the participants and the opinions of the participants were taken with focus group interviews. In the third part of the study the guide of the model and the data collection tools to be used in the model were developed in line with the opinions obtained from the second part.

The developed teacher evaluation model was aimed to ensure the professional development of teachers in the qualifications of planning and preparation, instruction,

reflective thinking, communication and collaboration, and professional responsibilities. Data aimed to be collected from multiple sources with the help of the developed data collection tools (classroom observation form, pre-observation, and post-observation interview forms, self-evaluation form, lesson plan review form, and professional responsibilities evaluation form). Functional reports covering the holistic evaluation of the teacher would be shared with the teacher at the end of each academic term, giving the teacher the right to speak through face-to-face meetings with the school principal.

Keywords: Teacher Evaluation Model, Participatory Teacher Evaluation, Professional Development, Case Study

ÖZ

KATILIMCI VE ÖĞRETMEN MESLEKİ GELİŞİMİNİ AMAÇLAYAN BİR ÖĞRETMEN DEĞERLENDİRME MODELİ GELİŞTİRME: VAKA ÇALIŞMASI

GÜNEŞ SAVUL, Emel

Doktora, Eğitim Bilimleri Bölümü

Tez Yöneticisi: Doç. Dr. YEŞİM ÇAPA-AYDIN

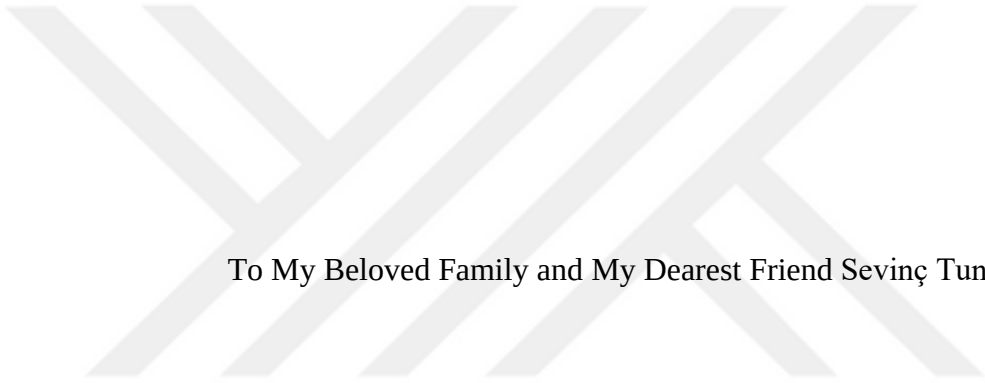
Ekim 2022, 395 pages

Bu örnek olay çalışması ile, öğretmenlerin mesleki gelişimini desteklemek için katılımcı bir öğretmen değerlendirme modeli geliştirilmesi amaçlanmıştır. Bu amaçla bir özel ilköğretim okulu belirlenmiş ve çalışma 2018-2020 eğitim öğretim yılında gerçekleştirilmiştir. Katılımcıları bu okulda görev yapan sınıf öğretmenleri, müdürler ve uzmanlar oluşturmuştur.

Araştırmanın ilk bölümü, katılımcıların mevcut öğretmen değerlendirme ve mesleki gelişim uygulamalarına ilişkin algılarının derinlemesine bir analizini sağlamayı amaçlamıştır. Araştırmanın ikinci bölümünde araştırmacı, birinci bölümden elde ettiği görüşler doğrultusunda bir model tasarlamıştır. Katılımcılara tasarlanmış model tanıtılmış ve odak grup görüşmeleri ile katılımcıların görüşleri alınmıştır. Çalışmanın üçüncü bölümünde, ikinci bölümden elde edilen görüşler doğrultusunda modelin kılavuzu ve modelde kullanılacak veri toplama araçları geliştirilmiştir.

Geliştirilen öğretmen değerlendirme modeli ile öğretmenlerin planlama ve hazırlık, öğretim, yansıtıcı düşünme, iletişim ve iş birliği ve mesleki sorumluluk alanlarında gelişimlerinin sağlanması amaçlanmıştır. Verilerin modelde kullanılmak üzere geliştirilen veri toplama araçları (sınıf gözlem formu, ön gözlem ve gözlem sonrası görüşme formları, öz değerlendirme formu, ders planı gözden geçirme formu ve mesleki sorumluluk değerlendirme formu) ile toplanması planlanmıştır. Öğretmenin bütüncül değerlendirmesini kapsayan mesleki gelişim raporlarının öğretmenlerle her akademik dönem sonunda, okul müdürü ile yapılan yüz yüze görüşmelerle paylaşılması ve bu görüşmelere öğretmene söz hakkı verilmesi planlanmıştır.

Anahtar Kelimeler: Öğretmen Değerlendirme Modeli, Katılımcı Öğretmen Değerlendirmesi, Mesleki Gelişim, Vaka Çalışması



To My Beloved Family and My Dearest Friend Sevinç Tunalı

ACKNOWLEDGEMENTS

First and foremost, I thank my esteemed supervisor Assoc. Prof. Dr. Yeřim apa Aydın for her support, motivation, guidance, and encouragement throughout all phases of this dissertation. I want to thank my thesis supervising committee members Prof. Dr. Fatma Hazır Bıkmaz and Prof. Dr. Cennet Engin in helping me to improve my dissertation and always ensuring that I was moving in the right direction. I would also like to thank my dissertation examining committee members Prof. Dr. Soner Yıldırım and Assoc. Prof. Dr. Berna Aslan for their valuable feedback and comments.

I am extremely grateful for the contribution and inspiration of my dear husband Gökhan Savul, who supported me during this thesis process that I have been trying to write since the day we got married, and who endured all my whinings during this process. I would like to thank for your believing in me that I will successfully complete this thesis from the very beginning to the end.

I would like to thank my dear daughter İdil Deniz for giving me joy and energy with her presence every day and for helping me to write my thesis at every opportunity, both during my pregnancy and after my birth.

To my mother Aliye Güneř and my father Mehmet Güneř, who, in addition to their duties as grandparents, provided me with all kinds of support without hesitation.

I would also like to thank my dear friend Ömer Ergem, who gifted me his own computer when my computer was on the verge of breaking down.

For this doctoral thesis, I would like to thank my dear friend Sevin Tunalı, whom I lost two years ago, and even dedicate this thesis to her. My dear friend, who sends resources for my thesis from every library she visits, is ready to help at every opportunity, motivates me to work at any time, and is always by my side, if you were alive, we would be celebrating with you right now. I'm sure you hear me somewhere. I love you so much and miss you every day

TABLE OF CONTENTS

PLAGIARISM	iii
ABSTRACT.....	iv
ÖZ	vi
DEDICATION.....	viii
ACKNOWLEDGEMENTS.....	ix
TABLE OF CONTENTS.....	x
LIST OF TABLES.....	xv
LIST OF FIGURES	xvii
LIST OF ABBREVIATIONS.....	xviii
CHAPTERS	
1. INTRODUCTION	1
1.1 Background to the Study	1
1.2 Purpose of the Study	8
1.3 Research Questions	9
1.4 Significance of the Study	10
1.5 Definition of Terms.....	15
2. LITERATURE REVIEW	16
2.1 History and Foundations of Teacher Evaluation.....	16
2.1.1 Development of Teacher Evaluation Processes in the USA.....	17
2.1.1.1 The Process of Changing Policies and Proposed Reforms	17
2.1.1.2 Pioneering Models Used in Teacher Evaluation.....	25
2.1.2 Teacher Evaluation Processes in Other Countries.....	30
2.1.3 The Development of Teacher Evaluation Process in Turkey	32
2.2 Components and Characteristics of Effective Teacher Evaluation.....	39
2.2.1 Role of the Teacher, Teacher Competency and Performance Standards ..	40
2.2.2 Purpose of Teacher Evaluation.....	42

2.2.2.1 The Accountability Function of Evaluation Model	43
2.2.2.2 Improving Teacher's Professional Development	45
2.2.3 Evaluator	46
2.2.4 Data Collection for Teacher Evaluation	50
2.2.4.1 Classroom Observation	50
2.2.4.2 Individual Interviews and Debriefing with Observation.....	52
2.2.4.3 Surveys/Questionnaires.....	53
2.2.4.4 Self-Assessment, Self-Reflection, and Self-Study.....	55
2.2.4.5 Planning Documents, Teaching Artifacts, and Portfolios.....	56
2.2.4.6 Student Achievement Scores-Value Added Measures.....	57
2.2.5 Use of Evaluation Results and Feedback Process	60
2.2.7 Credible and Trustworthy Environment	64
2.3 Teachers Active Participation in Evaluation Process.....	66
2.4 A School Specific Teacher Evaluation.....	67
2.5 Research Studies on Teacher Evaluation	69
2.5.1 Research Studies on Teacher Evaluation and Participatory Approach.....	69
2.5.2 Research Studies on Problems in Teacher Evaluation.....	71
2.5.3 Research Studies on Effective Features of Teacher Evaluation	72
2.5.4 Research Studies on Developing Teacher Evaluation Models	78
2.6 Summary of Literature Review	79
3. METHOD.....	86
3.1 Research Design	86
3.2 Research Questions	89
3.3 Context of The Study	90
3.4 Data Sources.....	92
3.4.1 Features of the Case	92
3.4.2 Selection of the Participants	93
3.5 Data Collection Instruments	95
3.5.1 Semi-Structured Interview Forms.....	96
3.5.2 Focus Group Interview Form.....	99
3.6 Data Collection Procedures	100
3.6.1 One-to-one Interview.....	101

3.6.2 Focus Group Interview	103
3.6.3 Online Feedback Forms.....	104
3.6.4 Pilot Testing of Teacher Evaluation Tools	104
3.7 Role of the Researcher	105
3.8 Data Analysis	106
3.8.1 Preparation for the Data Analysis Process	107
3.8.2 Coding and Constructing Categories	107
3.8.3 Reporting of the Findings	110
3.9 Trustworthiness of the Study	110
3.9.1 Credibility	110
3.9.2 Transferability	112
3.9.3 Dependability.....	112
3.9.4 Confirmability	113
3.10 Limitations of the Study	114
4. FINDINGS.....	116
4.1 First Part of The Study	116
4.1.1 Teacher Evaluation Practices Carried Out (Research Question 1).....	117
4.1.1.1 Evaluation Process	117
4.1.1.2 Use of Evaluation Data	121
4.1.2 Perceived Strength and Weaknesses (Research Question 2).....	122
4.1.2.1 Propriety Problems.....	122
4.1.2.2 Utility Problems	125
4.1.2.3 Accuracy Problems	129
4.1.2.3.1 Misinterpretation of Teacher Effectiveness	129
4.1.2.3.2 Bias identification	131
4.1.2.3.3 Valid Judgements	134
4.1.3 Recommendations for Developing Effective Teacher Evaluation Model (Research Question 3)	137
4.1.3.1 Recommendations for Evaluation Process.....	137
4.1.3.1.1 Recommendations for the Purpose of Evaluation and Qualifications	137
4.1.3.1.2 Recommendations for Evaluation Schedule	147
4.1.3.1.3. Recommendations for Evaluator Features	148

4.1.3.1.4 Recommendations for Enabling Evaluation Dynamics.....	153
4.1.3.1.5 Recommendations for the Evaluator	154
4.1.3.1.6 Recommendations for Evaluation Method.....	159
4.1.3.2 Recommendations for the Use of Evaluation Data	165
4.1.3.2.1 Access to Evaluation Information	165
4.1.3.2.2. Recommendations for Propriety	166
4.1.3.2.3 Recommendations for Utility	168
4.1.3.2.4 Recommendations for Accuracy	170
4.1.4 Participant Opinions for Professional Development Processes (Research Question 4).....	175
4.1.4.1 Planning the Process of Professional Development.....	175
4.1.4.2 Perceived Strength and Weaknesses of Professional Development Practices	177
4.1.4.3. Needs for Professional Development Practices	182
4.2 Second Part of The Study	184
4.2.1 The Initial Participatory Teacher Evaluation Model	185
4.2.2 The Opinions of The Participants Regarding the Initial Participatory Teacher Evaluation Model (Research Question 5).....	188
4.2.2.1 Important Issues Planned to be Included in the model	188
4.2.2.2 Organization of the Qualifications	193
4.2.2.3 Schedule of the Model	198
4.2.2.4 Access to Evaluation Information and Providing Propriety	204
4.2.2.5 Feasibility of the Model	207
4.2.2.6 The Guide and Tools of the Model	209
4.3 Third Part of The Study	214
4.3.1 The Guide and The Tools of The Teacher Evaluation Model	214
4.3.2 Opinions of the Participants Regarding the Guide of The Model (Research Question 6).....	220
4.3.3 Validity Evidence of Teacher Evaluation Tools (Research Question 7).....	221
4.3.3.1 Feedback About the General Structure of The Tools.....	223
4.3.3.2 Lesson Plan Evaluation Form	224
4.3.3.3 Pre-Observation Interview Form, Observation Form, and Post-Observation Interview Form	225
4.3.3.4 Communication and Cooperation Evaluation Form	227

4.3.3.5. Professional Responsibility Evaluation Form.....	228
4.4 Summary of Findings	229
5. DISCUSSION AND IMPLICATIONS	234
5.1 Current Practices and Problems About Teacher Evaluation	234
5.2 Teacher Evaluation Model Needed	239
5.3 A School-Based, Participatory Teacher Evaluation Model	259
5.4 Implications for Educational Practices.....	262
5.5 Implications for Further Research.....	265
REFERENCES	268
APPENDICES	
A. TEACHER INTERVIEW SCHEDULE	300
B. PRINCIPALS INTERVIEW SCHEDULE	304
C. EXPERT INTERVIEW SCHEDULE.....	309
D. FOCUS GROUP INTERVIEW SCHEDULE	313
E. THE CODEBOOK OF 1 ST DATA COLLECTION PROCESS (INDIVIDUAL INTERVIEWS).....	316
F. THE CODEBOOK OF 2 ND DATA COLLECTION PROCESS (FOCUS GROUP INTERVIEWS).....	325
G. THE INITIAL TEACHER EVALUATION MODEL.....	328
H. PART OF THE MODEL’S GUIDE	331
I. EXAMPLES OF DATA COLLECTION TOOLS OF THE MODEL	340
J. APPROVAL OF THE METU HUMAN SUBJECTS ETHICS COMMITTEE..	342
K. APPROVAL OF THE PRIVATE INSTITUTION (CASE).....	343
L. CURRICULUM VITAE	344
M. TURKISH SUMMARY / TÜRKÇE ÖZET.....	345
N. TEZ İZİN FORMU / THESIS PERMISSION FORM	375

LIST OF TABLES

Table 3.1 The Features of the Case	93
Table 3.2 Teachers' Grade Levels and Years of Experience	94
Table 3.3 Focus Group Participants	95
Table 3.4 Part of Participant T1's Responses, Codes, and Categories	109
Table 4.1 Codes for The Evaluation Process	117
Table 4.2 Codes for The Use of Evaluation Data	121
Table 4.3 Codes for Propriety Problems.....	122
Table 4.4 Codes for Utility Problems	125
Table 4.5 Codes for Accuracy Problems.....	129
Table 4.6 Codes for The Purpose of The Evaluation Model and the Qualification Areas	138
Table 4.7 Codes for The Evaluation Schedule.....	147
Table 4.8 Codes for The Evaluator Features.....	148
Table 4.9 Codes for The Evaluation Dynamics	153
Table 4.10 Codes for The Evaluator	154
Table 4.11 Codes for The Evaluation Method	159
Table 4.12 Codes for The Access to Evaluation Results	165
Table 4.13 Codes for Providing Propriety	166
Table 4.14 Codes for Providing Utility.....	168
Table 4.15 Codes for Providing Accuracy.....	171
Table 4.16 Codes for The Planning Process of Professional Development.....	175
Table 4.17 Codes for The Strengths and Problems of Professional Development Processes	178
Table 4.18 Codes for The Professional Development Needs	182
Table 4.19 Codes and Themes for The Important Issues in The Initial Model	188

Table 4.20 Codes and Themes for The Qualifications to be Evaluated in The Initial Model	193
Table 4.21 Codes and Themes for The Schedule of The Initial Model	199
Table 4.22 Codes and Themes for Accessing Evaluation Information and Providing Propriety.....	204
Table 4.23 Codes and Themes for the Feasibility of the Model	207
Table 4.24 Codes and Themes for the Guide and The Tools of the Model	209
Table 4.25 Tools, Evaluators, and Validation Process	223



LIST OF FIGURES

Figure 3.1 The overall design of the study.....	88
Figure 4.1 Teacher Qualification Areas and Subdimensions Included in The Initial Model	187
Figure 4.2 Teacher Qualification Areas and Subdimensions Included in The Guide of the Model	216

LIST OF ABBREVIATIONS

CDE	Centre for Development and Enterprise
ESSA	Every Student Succeeds Act
MET	Measures of Effective Teaching
NCTQ	National Council on Teacher Quality
NCLB	No Child Left Behind
NEA	National Education Association
NIET	National Institute for Excellence in Teaching
OECD	Organization for Economic Co-operation and Development
OFSTED	Office for Standards in Education
TALIS	Teaching and Learning International Survey
TNTP	The New Teacher Project
VAM	Value Added Models

CHAPTER 1

INTRODUCTION

This chapter includes the background information, significance, and purpose of the study. Furthermore, research questions and definitions of key terms are also provided in this chapter.

1.1 Background to the Study

The core task of education is to ensure that students learn through school experience and that students gain their current and future selves, so teachers' learning directly affects their students. (Netolicky, 2020). An education system is effective when teachers improve student performance and develop the full potential of each student. In other words, teachers stay at the center of educational systems, and they are reviewed as the most influential school-related force for student achievement (Darling-Hammond, 2014; Fullan, 2011; Papay, 2012; Wiliam, 2018). Seen through this lens, teachers matter, and the work teachers do in the classroom also matters (Organization for Economic Co-operation and Development [OECD], 2018; Stronge, 2006; Weisberg et al., 2009). Effective teaching has been defined in different ways throughout history. In the past, the teacher was regarded as the gatekeepers of knowledge, and the effective teacher provided the student's access to knowledge in a controlled manner. Today, both the research on how the brain learns and the developments in the field of technology show that the teacher-student interaction is important in learning and teacher who raises students that can easily access information, learn with their own effort, and experience, and have high-level thinking skills such as critical thinking and problem solving, is an effective teacher. Although effective teaching has brought many different definitions over the years, it can be said

that the qualifications areas that an effective teacher should have determined together with crucial contributions to students' learning and achievement (Centre for Development and Enterprise [CDE], 2015; Clotfelter et al., 2007; Darling-Hammond, 2006; Hanushek, 1992; Kim & Sun, 2021). The word "crucial contribution" here is also important because the teacher influences the student not only on the current learning process, but also on the future enjoyment of learning. As Tucker & Stronge (2005) mentioned, when students are taught by teachers with a high-performance level during a year, they continue to take advantage of this good teaching in the coming years, on the contrary, the teaching provided by a low-performing teacher negatively affects the student's learning in the coming years. Besides, it cannot be thought that the effectiveness of the teacher is valid only in the classroom environment. The influence of the teacher goes beyond the walls of the classroom, and teachers also prepare students step by step for the situations that await them in the future. Teachers also play an essential role in ensuring human power which is also an indispensable element of today's competitive world and worldwide societies are heavily invested in the teaching profession (Organization for Economic Co-operation and Development [OECD], 2019; Wiliam, 2018).

The quality of any teaching process that takes place could be as effective as the quality of the teacher and improving teacher quality is one of the most promising and essential strategies to enhance the quality of teaching. The idea of improving the teacher's quality and therefore the quality of teaching has become a significant focus for policymakers and is considered a primary tool to be improved (CDE, 2015; Darling-Hammond, 2010; Darling-Hammond, 2014; Stronge, 2006; Şahan, 2011). Since teacher quality has been identified as the most important factor affecting student achievement, emphasis has been placed on promoting teacher quality through teacher evaluation systems (Darling-Hammond, 2012; Looney, 2011; Papay, 2012). Teacher improvement is possible through teacher evaluation because something that is not evaluated cannot be changed or developed. In general teacher evaluation can be defined as a systematic process used to review teachers' performance both in school and classroom to provide constructive feedback and guidance for professional

development (Ornstein, 1985; Danielson, 2011; Darling-Hammond, 2010). The definition of teacher evaluation also shows us the importance of these systems, but the teacher evaluation system is not only important but necessary (Özbek & Taneri, 2019). Policymakers and scholars are identifying the teacher evaluation systems as the crucial part of improving teachers' quality to raise the students' academic performance (Gordon, et al., 2006).

Teacher evaluation has a long history and many attempts that take their roots from past educational reforms. The magnitude of the teacher's influence on student achievement has led to a continuous change in teacher evaluation systems. Therefore, high budgets have been allocated to teacher evaluation reforms (Dee et al., 2021). The changing education policies and the reforms (e.g., Race to the Top, Teacher Incentive Fund, the Measures of Effective Teaching Project, No Child Left Behind waivers, and Intensive Partnerships for Effective Teaching) put forward have caused the teacher evaluation systems applied in schools to be updated or restructured (Dee et al., 2021; Donaldson & Woulfin, 2018; Garrett & Steinberg, 2015). As in the United States, reforms proposed by every government have led to the design and implementation of new teacher models in most states and school districts (Steinberg & Donaldson, 2016). Recently most districts and schools are implementing intensive and high-stake (threat of dismissal based on an evaluation) teacher evaluation systems, including classroom observations, supervisory conferences, test scores of students (Donaldson & Woulfin, 2018). Reforms or implications conducted to develop educational system and for more effective teacher evaluation models were not as effective as expected. Besides teacher evaluation reform as a system that is still in progress and many states struggled to implement a teacher evaluation system (McGuinn, 2012). One of the most important reasons why this effect is not at the expected level is giving more weight to normative measures such as value-added scores rather than criteria-reference measures (Kraft & Gilmour, 2017). While the incompetence in field of evaluation and lack of training of the people who made the evaluations reduced the impact of these reform. Furthermore, making evaluations in order to determine the failure of the teacher and to impose sanctions on the teacher caused the reforms to move away from the aims aimed at

work. On the other hand, in 2016 “Every Student Succeeds Act (ESSA)”, signaled states can have more freedom in changing teacher evaluation policies and states are given local control over teacher evaluation. States also allowed districts to develop and implement more specific teacher evaluation systems. Thus, the states began to look for ways to alleviate the heavy burden of the Value-Added Models. Many states today offer more alternatives to teacher evaluation by measuring the relationships between student achievement and teacher effectiveness and state teacher evaluation plans also include more formative teacher feedback (Close et al., 2020).

Research and reports on OECD countries help to have an idea about teacher evaluation models in these countries. According to the “Teachers for the 21st Century Using Evaluation To improve Teaching” report, it was revealed that some OECD countries have relatively poor evaluation structure, teachers cannot benefit from any feedback given to improve their instruction, and evaluation processes do not take place in policy frameworks (Organization for Economic Co-operation and Development [OECD], 2013a). When we look at the process in Turkey, it can be said that the foundations of teacher evaluation go back to the first years of the Republic. To the recent past from those years, the teacher evaluation initiatives in Turkey were carried out with external evaluation mostly, with the evaluation of the inspector appointed by the Ministry of Education (Education Research and Development Department, 1995; Pehlivan et al., 2001). In the following years, many attempts were carried out to determine effective teacher qualifications and the behaviors to be measured. Over time, inspectors were prevented from making classroom observations by limiting the external evaluations because their competencies might not be suitable for evaluation (Buluç, 1997). However, the inability to replace the inspectors with anyone who will observe the teaching in the classroom and the failure to propose a model that can systematically evaluate teacher performance has caused the teacher evaluation studies in Turkey to lose their effectiveness with each passing year. Furthermore, OECD reported that in Turkey, a quantitative approach is used for teacher appraisals meaning that teachers do not always receive qualitative feedback to guide improvement (Organization for Economic Co-operation and Development [OECD], 2020b). Both teacher evaluations

and school supervisions are not yet fully used to support the development of the teacher and it is crucial to build capacity among school teams to ensure that teacher evaluation, school inspection, and systems evaluation are aligned with national priorities that result in progress (Organization for Economic Co-operation and Development [OECD], 2020a).

Recognizing the importance of teacher evaluation, reforms, practices, models, and initiatives have been proposed by different countries and districts of United States contributed to improve the teacher evaluation process. Perhaps the most important contribution of the practices made in line with these suggestions is that they help determine some basic components that should be in an effective teacher evaluation model. Although it is not possible to talk about a single and effective teacher evaluation model, establishing detailed standards or criteria by which teacher efficacy is clearly defined, conducting standards-based teacher observations throughout the year and using multiple measures to collect evidence credible teacher evaluation systems can be developed (Darling-Hammond et al., 2012; Weisberg et al., 2009, Putman et al., 2018). Teaching is a highly complex, interactive, structured process that is unlikely to result in the absolute success of students, and a performance assessment based on multiple data sources may yield more objective results when assessing the performance of teachers undertaking such a multidimensional task (Kahya & Hoşgörür, 2020). Besides, it is also possible to say that the purpose of successful and strong teacher evaluation models is tied with the professional development of the teacher to identify areas in which a teacher is performing well and areas which a teacher needs to improve (Putman et al., 2018). All in all, it was evident that the evaluation process, which is based on the development of the teacher, provides timely formative feedback, offers teachers to participate in the evaluation process, conducts data collection procedures by collecting evidence from more than one source is more effective and stronger (Danielson, 2013; Darling-Hammond, 2012; Derrington & Brandon, 2019; Marzano, 2012; OECD, 2013a).

Establishing the purpose of the model and developing the design of the model in accordance with this purpose takes part at the beginning of these important features.

Teacher evaluation can have both individual and organizational levels and two purposes: accountability and improvement (Hopkins et al., 2016). The view of predicting teacher success with the help of the scores of students' success obtained from standardized tests, which is mostly carried out for the purpose of accountability, also known as value-added systems, has remained popular for many years. These systems do not support the evaluation of teachers who do not evaluate with standardized tests (Toch & Rothman, 2008). Moreover, since multiple-choice tests are only used to measure low-level thinking skills, ignoring how the teacher gains high-level thinking skills has caused these systems to lose their popularity. The use of such measurements for high-stake decision-making, such as teacher termination, merit pay, and denial of teacher tenure also raises serious questioning of the value-added system's purported usefulness (Amrein-Beardsley & Geiger, 2017). To evaluate teachers fairly, accurately, and credibly the aim of evaluation systems should be linked with professional development not using the poor performance evidence to dismiss teacher (Weisberg et al., 2009). Teacher evaluations should connect to growth and development, and it is essential to design evaluation systems to provide opportunities for professional learning for teachers (Danielson & McGreal, 2000; Darling-Hammond et al., 2017; Tarhan et al., 2019; Weisberg et al., 2009).

It has been known for many years that it is necessary to evaluate the teacher in order to improve, but the important question we need to ask here is what kind of evaluation will make effective and fair judgments. Undoubtedly, it is an indispensable and crucial step to determine the qualifications to be developed in ensuring the development of the teacher through evaluation. Moreover, if an evaluation is made to improve the qualifications of the teacher, it is necessary to examine in detail which qualification areas will be evaluated and to include explanatory criteria in order for the decisions to be taken by this evaluation to be effective, fair, and acceptable or justifiable by the evaluated person. Every teacher's needs for professional development are different, and teachers should be provided with ongoing opportunities to learn new skills and continually improve themselves in the profession by setting standards that reflect the features of effective teaching to guide evaluation (Darling-Hammond et al., 2017;

Derrington & Brandon, 2019). Teacher evaluation systems will not be as effective if the evaluation criteria used miss important components of teaching qualifications that could aid teacher development or more effective staff practices (Master, 2014). Undoubtedly, it is an indispensable and crucial step to determine the qualifications to be developed in ensuring the development of the teacher through evaluation.

Another critical element of an evaluation that supports the professional development of the teacher is conducting the effective feedback process (Kane et al., 2014; Organization for Economic Co-operation and Development [OECD], 2009b). It is essential to distinguish teachers who are succeeding and struggling; furthermore, an effective teacher evaluation model should provide feedback and continuous improvement (Darling-Hammond, 2014). The feedback process should include constructive approaches, conducted with valid tools including specific criteria, and should be given within a sincere and professional dialogue (Gordon & McGhee, 2019; Danielson & McGreal, 2000; Lynda et al., 2021). Teacher evaluation that combines evaluation results with productive feedback for professional development can increase teacher effectiveness. Many international studies have suggested that evaluation systems will increase teacher effectiveness if multiple indicators of teacher performance are used, designed to provide feedback to the teacher, and if the teacher were provided rich learning environment (CDE, 2015; Organization for Economic Co-operation and Development [OECD], 2013a). A landmark study of professional development found that, commonly, teachers do not get clear information about how to improve, nor do they feel that their professional development is customized based on their needs (Putman et al., 2018).

It is possible to talk about the effectiveness of teacher evaluation systems or models whose purpose is well defined, the qualification areas that will be measured are determined in detail and correctly, and at the end of the evaluation, effective feedback is given to the teacher to improve himself. On the other hand, ensuring the participation of the teacher in the development and implementation of teacher evaluation models or systems is very important for effectiveness. It can be said that there is an inconsistency between the classroom practices that teachers believe and

prefer to use and those required by the policies (Bonner et al., 2018; Muskin, 2015). For instance, many teachers working in different states of America are concerned about the accountability systems and value-added models imposed on them and do not believe in their effectiveness. Teachers are conveying that they have negative emotions and experiences related to accountability focused evaluations which also harmed their identities (Guenther, 2021). The creation of systems that support the teacher's autonomy and include the teacher's own expressions is a prerequisite for achieving the constantly renewed teaching goals (Bonner et al., 2018). In fact, the active participation of the teacher in the design of the evaluation process means sharing the responsibility. Shared responsibility is key to achieving both accountability and teacher professional development goals (Ryu, 2020). Shared responsibility and a collaborative school environment create more opportunities for teachers to build their professional capacity by allowing them to participate in school initiatives and decision-making. The supportive and participatory teacher evaluation process leads to teacher leadership, and teachers who participate in teacher evaluation become even more productive in later years (Taylor & Tyler, 2012). Asking teachers to contribute to the development of their own evaluation systems with their active participation in every stage of these systems make the developed evaluation models more effective (Danielson & McGreal, 2000; Marzano & Toth, 2013). Policy makers need to look for ways to involve teachers in developing and improving their assessment systems (Close & Amrein-Beardsley, 2018). To employ more effective teacher evaluation models in the future, it is very important to use more than one measure in evaluating teachers, to design teacher evaluation systems that emphasize formative uses, and to involve teachers throughout the process of creating and improving these systems (Close & Amrein-Beardsley, 2018).

1.2 Purpose of the Study

The purpose of this research was to plan, organize and design a participatory teacher evaluation model to promote teachers' professional development.

1.3 Research Questions

This research is composed of three parts. The first part of the research aimed to provide an in-depth analysis of perceptions of classroom teachers, principals, and experts on current teacher evaluation and professional development practices at the private school and to examine the suggestions of the participants in depth. More specifically, the research questions were:

1. What kind of teacher evaluation practices are carried out for classroom teachers at this private school?
2. How are teacher evaluation practices perceived in terms of strengths and weaknesses by classroom teachers, principals, and experts at this private school?
3. What are the recommendations of this school's classroom teachers, principals, and experts to develop an effective teacher evaluation model?
4. What kind of professional development processes are carried out at this private school, and what is needed?

The second and third part of this research aimed to develop a participatory teacher evaluation model to promote teachers' professional development with data collection tools. More specifically, the research questions were:

5. What are the opinions of the participants regarding the initial teacher evaluation model?
6. What are the opinions of the participants regarding the guide of the model?
7. What are the validity evidence of teacher evaluation tools?

1.4 Significance of the Study

Effects of well-prepared teachers on student achievement can be stronger than the influences of student background factors such as poverty, language background, and minority status (Darling-Hammond, 2000). Unfortunately, although we believe that all students can learn to high standards, we still made little progress in answering why poor instruction is still being implemented in our schools. More importantly, an evaluation system to assess the instructional performance of teachers accurately to support who are responsible for this poor instruction in schools or to reward the excellence of hardworking teachers is still not fully developed (Grissom & Youngs, 2015; Kim & Sun, 2021; Weisberg et al., 2009). According to Bill and Melinda Gates Foundation research paper on the Measures of Effective Teaching (MET) project, failures in giving high-quality professional feedback to teachers is a threat to well-qualified instruction (Kane & Staiger, 2012).

Furthermore, evaluations based on scant evidence, relied on imprecise instruments, provided useless feedback to teachers, conducted by not adequately trained evaluators who are mostly lacking in expertise, and used performance ratings that are not considered for basic personnel decisions (Jerald, 2009). The inadequacy of the measurement tools, the limited data sources, and the use of student test scores as the only data source may cause an inappropriate evaluation of teacher effectiveness (Popham, 2013). Many teacher evaluation models mainly rely on poor predictors such as paper-pencil tests of essential academic skills and subject matter of knowledge or classroom observations done by principals who can define effective teaching in a limited way (Darling-Hammond, 2010). Teachers feel that evaluations do not reflect the quality of their work since the value-added measures in which statistical calculations are used for understanding teacher impact on students' achievement are thought to be biased and misleading. Besides, the classroom observation process has been rolled out due to the insufficient training of raters and problems in ensuring the fidelity of the instruments and procedures (Grissom & Youngs, 2015).

Although evaluating teachers and teacher evaluation reforms is a respectable and important issue in almost all countries, current teacher evaluation systems do little to help for teacher improvement and barely support personal decision-making processes (Darling-Hammond et al., 2012; National Council on Teacher Quality [NCTQ], 2011; National Education Association [NEA], 2012). In the “Rush to Judgment” report, it also revealed that the evaluation systems which emphasize teacher credentials more than teacher performance, assess teachers by a simple checklist item, and observe teacher behaviors that do not focus on the quality of the instruction has little value in improving the instruction (Toch & Rothman, 2008).

In 2009, the Teaching and Learning International Survey (TALIS) was conducted to collect data on related issues, including the recognition of teachers and rewards that teacher receive. The TALIS report showed that teachers from TALIS countries noted that appraisal and feedback are beneficial for their development. Still, several countries find it challenging to ensure that all teachers systematically receive effective appraisal and feedback from their employers (Organization for Economic Co-operation and Development [OECD], 2009a). Furthermore, report also pointed out that the current teacher evaluation systems lack the necessary support and incentives for teachers’ growth and also for the education teachers provide to students. More importantly, criticisms of teachers are often expressed not based on data but instead based on various generalizations done through personal observations and experiences. Teachers have little trust in teacher evaluation systems; in other words, they do not trust the scores gathered through evaluation because they think there is favoritism in the given scores. Besides, they also thought that provided feedback is idiosyncratic and not remarkable due to the evaluations made without setting clear criteria and their performance only evaluated by administrators using one-way communication (Danielson & McGreal, 2000). Teachers also perceive teacher evaluations as technical than functioning to promote teachers’ growth and development (McClure, 2008). It can be suggested that performance evaluation should be carried out not with an evaluation approach that will cause anxiety in the teacher, but by reducing the factors that negatively affect the perception of justice, that is, with an individualized

supervision system that will support the development of the teacher, discover his talents and monitor his own performance (Boydak Özcan & Karagözoğlu, 2020).

On the other hand, studies reveal that teachers are open to new evaluation models and value models that promote teachers' professional development (Campbell, 2014; Ramage, 2012; Pässe, 2015). Furthermore, teachers believe that they can build trust and find opportunities for their growth and development if evaluation models are adequately implemented by whom well equipped with necessary skills, if teachers are educated about the evaluation process, and if the evaluation is done in an environment in which teachers trusts the evaluators and feel collegiality (Campbell, 2014; Rucinski & Diersing, 2014; Walker, 2014). Rigorous teacher evaluation systems work well in promoting the improvement of the school if the system is designed and implemented carefully (Grissom & Youngs, 2015).

However, teacher evaluation reform is in its infancy and will go through many changes and iterations to reach maturity. Educational systems still need newly proposed teacher evaluation systems and models (Marzano & Toth, 2013). There is a need for a transformation for teacher evaluation models, which includes standard-based evaluations with explicit criteria, enables measuring the skill accurately and improve teachers' performance to develop effective teachers who can teach in powerful ways. For this transformation, teachers' abilities and capacity should be understood and assessed validly (Carbaugh et al., 2020; Darling-Hammond, 2010; Marzano, 2017). Instead of identifying different levels of teachers' performance by using a series of checklists, there is a need for schools to distinguish great from good, good from fair, and fair from poor (Danielson & McGreal, 2000; Kraft & Gilmour, 2017; Weisberg et al., 2009).

Teacher evaluation implementations and practices have been an issue for a long time and still calls for a reform affecting educators, policymakers, school systems, teachers, etc. (Reddy et al., 2016). Still, many countries are trying to reform their teacher evaluation models and go beyond value added models; furthermore, most states in the United States are overhauling the evaluation system for teachers and administrators

(Darling-Hammond, 2014). Evaluating teachers is generally seen as perfunctory and episodic events; therefore, dissatisfied evaluation models created new accountability requirements and transformed teacher evaluation models (Derrington, 2011). Studies on teacher evaluation were accelerated greatly in the last ten years, and international studies and practices on teacher evaluation already produced remarkable findings for those countries. Still, Turkey is not yet included in the performance evaluation (Özbek & Taneri, 2019). Evaluating teachers in the proper manner always had been an issue for Turkey and other countries. It is seen that this system is far from improving the teaching process in Turkey, since educational supervision is control-oriented and formal evaluations are limited (Memduhoğlu, 2012). Studies conducted in Turkey revealed that a new performance assessment model for the teacher is needed (Alay, 2006; Çavuş, 2010; Pehlivan et al., 2001). When the Turkish perspective is considered, researchers conducted studies to analyze the current situation and take teachers, principals, and supervisors' opinions. However, there is still needed to propose a model for teacher evaluation and their professional development.

The importance of teacher evaluation models has been on the agenda for a long time, and many studies have been conducted on the effectiveness of these models. Considering why models or systems are still not successfully implemented, and new models are constantly needed, evaluation of teachers as “effective” at a high rate, far from objectivity, seems like the main problem in teacher evaluation practices (Darling-Hammond, 2014; Popham, 2013). One of the reasons teacher evaluation models are still not as effective as desired is to examine the effectiveness of the models prepared within the scope of the education policy; in other words, the models are recommended from an external source (Kim & Sun, 2021). To examine the effectiveness of the implementation of education policies, it is necessary to examine the ability of the education system and the willingness of the actors to change and engage (Viennet & Pont, 2017). Therefore, evaluation and its impact must be viewed from teachers' perspective (Derrington & Brandon, 2019). The participation of stakeholders in designing, developing, and implementing evaluation systems increases the system's applicability (Danielson & McGreal, 2000; Howard & Gullickson, 2009). Teachers,

who have a say in their development, believe that the evaluation system being applied is entirely for their professional development and the results only are used for their benefit (Attard, 2016; Brandenburg et al., 2017). Teacher assessment systems need to be evaluated for effectiveness and include teacher perceptions in the creation and monitoring of these teacher assessment systems (Paufler et al., 2020).

Another reason for the failure of teacher evaluation models is the neglect of needs analysis used to determine the difficulties to support teachers in the improvement activities and the use of federal practices rather than local improvement (Dee & Dizon-Ross, 2019). Teacher evaluation processes have evolved over time from systems where teachers are monitored by external local authorities to ensure compliance, to a system where teachers and administrators work collaboratively using research-based methods to measure teacher performance (Fuller, 2022). Both in primary and secondary levels, a trend towards more decentralization and school autonomy is increasing in the evaluation and assessment of students, teachers, school leaders, and schools. An effective teacher evaluation should be conducive to a secure communication environment that allows teachers to discuss their problems with others and learn from them via collective participation of teachers from the same school (Desimone, 2009). A school-based assessment, which provides an opportunity to take a closer look at a teacher's classroom practices and provides more significant opportunities for communication and collaboration among teachers, is more conducive to establishing a development-oriented teacher assessment (Chen et al., 2021). School-based assessments create a good atmosphere for teacher development, make teachers natural learners in this environment, and make teachers more willing to learn by helping teachers gain commitment to the process (Kurum & Cinkir, 2019; Wong & Li, 2010). Although researchers have measured teacher perceptions of the evaluation process, limited research has been conducted on teachers' perceptions of the process in schools, creating a research gap that this study aims to address (Fuller, 2022).

All in all fair and effective teacher evaluation models, which providing timely and meaningful feedback to teachers, including observations done throughout multiple sources and training educators to become expert evaluators in order to evaluate

teachers effectively, are needed (Bill and Melinda Gates Foundation, 2010; Darling-Hammond et al., 2012; Darling-Hammond, 2014; Dee et al., 2021). Within the scope of this research, it is expected that the proposal of the teacher evaluation model, developed specifically for the structure and culture of a school, with the participation of all relevant stakeholders in the school and aiming to support professional development, would be effective in preventing the mentioned problems.

1.5. Definition of Terms

Teacher Evaluation: Systematic assessment procedure for collecting evidence from multiple sources about the qualification concerning a professional role of the teachers to provide constructive feedback, help reinforce outstanding service and provide the opportunity for development practices (Haefele, 1993; Howard & Gullickson, 2009).

Teacher Evaluation Model: They are systematic platforms created for educational institutions to evaluate teachers accurately for certain purposes.

Teacher Professional Development: Activities used to develop a teacher's knowledge, expertise, and other characters (OECD, 2009a).

Participatory Evaluation: Engaging a large number of potentially interested members of an organization and involving practice-based decision-makers who are usually organization members and primary users of evaluations to create support which is also called a stakeholder-based evaluation model (Cousins & Earl, 1995)

Department: It refers to a group of classroom teachers working at the same grade level

Head of Department: A teacher who is responsible for a group of classroom teachers who teach at the same level and are located above them as a teacher manager

CHAPTER 2

LITERATURE REVIEW

2.1 History and Foundations of Teacher Evaluation

Teacher evaluation is a subject that is as old as schooling itself, and this situation reveals that we should also pay attention to the historical development of what is expected of an effective teacher. Studies and trends regarding who an effective teacher is, and which international standards define an effective teacher are in constant development and transformation. These studies carried out in history also reveal the changing shift in the characteristics attributed to the effective teacher and their in-class practices. Furthermore, to understand the historical change in teacher evaluation, understanding the historical process of the teaching profession is vital because teacher evaluation systems that developed in the past reflect the dominant belief on teacher effectiveness at that time.

Although several schools have evaluated teachers from a formal perspective, it is not possible to talk about the formal evaluation of teachers until the 20th century (Shinkfield & Stufflebeam, 1997). Changes in the meaning of effective teaching bring changes in evaluation criteria. Teachers were considered gatekeepers of knowledge in earlier times, meaning that they possessed knowledge and controlled access to it. So, it has been an inevitable result that past strategies focused only on some behavior of teachers, like enabling students to sit quietly in their classrooms (Danielson & McGreal, 2000). Today, many studies on the brain and how it learns revealed that developing and evaluating students' high level of thinking skills like critical thinking and problem-solving is essential, so new approaches are needed for teaching (Danielson & McGreal, 2000). Furthermore, in contemporary society, learners bring

the knowledge through their own experiences, and access to knowledge is open to literate people who have access to libraries and the internet (Graham. & Berman, 2015). So that being a gatekeeper of knowledge is not a valid term for today's teachers anymore. For that reason, the historical process of teacher evaluation is given together with the change in the meaning attributed to effective teaching. The chronological process of teacher evaluation in the USA, OECD countries, and Turkey have been shared under the following headings.

2.1.1 Development of Teacher Evaluation Processes in the USA

While explaining the historical development of teacher evaluation processes in the USA in this section, first, proposed reforms in line with changing education policies and different perspectives on teacher evaluation are explained. Secondly, the leading teacher evaluation models that have been developed in the historical process were mentioned.

2.1.1.1 The Process of Changing Policies and Proposed Reforms

In the 1820 and 1860 Industrial Revolution, immigrant workers were attracted by urban development in the United States. In this period, access to education was limited, teachers lacked qualifications, textbook recitation remained the primary learning method, and most Americans were illiterate (Jewel, 2017). However, as the schooling rate increased and education became available to all, the curricula took a more academic structure, which has revealed the need for teachers who were better trained by a principal or expert (Marzano et al., 2011). The fact that teachers had to be trained brought that they need to be inspected. After 1840 teacher autonomy decreased, and administrative control increased, so that community leaders started to inspect teachers in classrooms and school districts (Jewel, 2017).

In the 1900s, there was a dominant understanding that determining some teacher characteristics could determine whether the teacher is good or bad. Kratz's research called "Characteristics of the Best Teacher as Recognized by Children" was considered a pioneer in teacher evaluation (Millman & Darling-Hammond, 1991). In

this study, 2411 students from 2nd grade to the 8th grade were asked to define the characteristics of their best teachers. As a result, most of the students (87%) stated that their best teachers were helpful. Still, another remarkable result was that they expressed the personal appearance (58%) of their teachers as important (Kratz, 1896). In the 1940s, studies like trade research were conducted to examine presage variables related to the teacher's character, such as voice, appearance, affective moderation, appearance, willingness. With these studies, teachers with these intended features are viewed as effective teaching (Danielson & McGreal, 2000). Through the studies conducted in this period, which continued their effect for a long time, it was revealed that teachers who spoke well, had a good appearance, were enthusiastic, and were confident was defined as good teachers. Still, in the middle of the 20th century, it was realized that it was unnecessary to relate the quality of teaching with the teacher's characteristics and understanding in teacher evaluation started to focus on the characteristics of teachers in promoting student learning (Shinkfield & Stufflebeam, 1997).

In the 1950s, when the students' cognitive learning became important and when intelligence tests took their place in history as a popular concept, it was tried to collect data about the characteristics and attitudes of the teachers based on survey research (LeCompte, 2009). On the other hand, in the 1950s, Domas and Tiedaman reviewed more than one thousand studies and revealed no relationship between the teacher's characteristics and student achievement (Briesch et al., 2018). By-product of Hawthorne applied research in social sciences conducted by Mayo revealed that human relationship is essential in learning. With the emergence of this idea, there were parallel changes in teacher evaluation. For instance, teacher-student interaction began to be noticed, and it was believed that change in the development of the teacher would affect the learning of the students as well (Shinkfield & Stufflebeam, 1997). In-class observations started with the curiosity to understand the behavior of the teacher in a classroom. The effect of the teacher on the student has been revealed and the studies have focused on results such as the negative effects of some teaching methods on the

student, ignoring the developmental levels of the students and not considering the classroom environment (Medley, 1979).

With the increasing importance of accountability in the 1960s and 1970s, teacher evaluation became increasingly important. During this period, teacher evaluations were systematically conducted in most of the schools in the USA, and these evaluations were made by the observations of school principals and other school administrators. In the 1960s and 1970s, correlational analysis was conducted to connect teacher enthusiasm and student achievement. Correlational research gained speed in this period to compensate for the deficiencies in previous research. In contrast, experimental and quasi-experimental designs in educational research began to rapidly take their place in the historical process (Gage, 1963). Teachers' ability of teaching the content especially in science and mathematics has become increasingly important, and students' success in these courses used for teacher evaluation (Danielson & McGreal, 2000). As it became essential to examine the teaching and document the teacher's behavior, observations were made in the classroom with studies using quantitative research methods (test scores and survey responses), and in-depth qualitative investigations alone began to be included (LeCompte, 2009). In the 1970s and 1980s, rating scales and checklists were introduced, but these scales and checklists were used only to access the summary of information by ignoring the process (Danielson & McGreal, 2000).

Although open to many misinterpretations, developments during 1960s and 1970s, confirmed the critical role that the teacher played in student learning. In 1982, Madeline Hunter and her colleagues at UCLA University, in recognition of the importance of motivation, retention, and transferring concepts, proposed the Hunter model to improve the teaching activities of the teacher in the classroom and to improve the decision-making behavior of the teacher (Stallings, 1985). This model, whose effects continue today, dominated the teaching views of the 1980s and started the education-oriented teacher development trend (Danielson & McGreal, 2000). Moreover, this model has contributed significantly to the education field, revealing the teacher's critical role in student learning.

In the 1980s and 1990s, teaching for understanding and use of knowledge became important. With the importance of understanding the complexity of learning as a cognitive process, the importance of teaching higher-level skills such as critical thinking, problem-solving, working in collaboration, lifelong learning had emerged. With this cognitive transformation, it was understood that the social learning structure, the basics of higher-level thinking, and how the student can make sense of knowledge needed to be discovered (Brophy, 1992). With these emerging concepts, the idea "the teacher is only the person presenting the information" has lost its importance, and how the teacher will bring these critical skills to the student has also become a topic of discussion. In the 1980s, as a result of "Teacher Evaluation: A Study of Effective Practices" study which conducted by the RAND group, it was revealed that teacher evaluations were not specific enough to improve teachers' pedagogical skills and teachers, who are also strongest proponents of a more precise and rigorous approach, also criticized the teacher evaluation processes (Wise et al., 1984). Inadequacies of teacher evaluation and the evidence to show inadequateness first appeared in that research, so that this research took its place in theoretical literature (Marzano & Toth, 2013). The dominant model of teacher evaluation was in trouble because the criteria used in the evaluation lacked validity due to being derived from vague notions and personal experience, and untrained evaluators were not able to make valid judgments (Haefele, 1993). In the 1980s and 1990s, concerns about the American economy, such as the need for students to have sufficient employment in the changing business world with the skills necessary to achieve success in business life, enabled making changes in the understanding of how teaching should be changed and how content should be taught (Danielson & McGreal, 2000). In the late 1980s and early 1990s, the emphasis on teacher evaluation began to shift from observation and teacher behavior to student achievement (Jewell, 2017). To measure student performance, performance measurement methods had to be established. New standards for performance were established across states, and the federal government provided financial grants to states that agreed to evaluate teachers based on student performance (Derrington & Brandon, 2019). Therefore, the teacher evaluation concept shifted from an inspection

model toward increased teacher observations based on standards to measure performance.

Although the word "quality", coined in 1996, was seen as an important new word in education at the time, it was perceived as positioning teachers as objects, and teachers still find the term dehumanizing today (Netolicky, 2020). With the coining of the term, teaching was treated as a list of competencies to be supervised, correlated with student results, and measured by standardized tests (Connell, 2009). To improve this list of competencies, during the decades 1998-2017, "holding teacher education accountable" emerged as a key approach to reforming teacher education in the United States (Cochran-Smith, 2019). Teaching began to be perceived as a profession that needed to be repaired, and teacher education started to be perceived as a process for recruitment, certification, or regulation.

Determining the teacher as the sole and most important person responsible for the student's learning and success brought the inclusion of the teacher and school in the accountability process. No Child Left Behind (NCLB) Act was proposed as the primary law in the United States from 2002 to 2015 for K-12 general education, which held schools responsible and accountable for students learning and achievement. To make practices competitive, states had to develop new teacher evaluation systems that used multiple performance measures. Public Schools of more than 30 states and the District of Columbia have changed their policies, including making data on student learning the important or most important factor in teacher evaluations (Doherty & Jacobs, 2013). NCLB is best described as a status model, meaning that it is used to reflect the percentage of students who were at specific levels of achievement. According to Marzano and Toth (2013), the reason for using this act as the status model was to leave excuses for student failure because students will learn and achieve, or rather progress especially in the lowest performing schools. The aim of this reform was holding schools, teachers, and students accountable for meeting higher standards, as measured by student performance on standard assessments. By achieving the aim, it was believed as administrators would better supervise public schools, teachers would teach better, and students would take their learning more seriously (Close et al.,

2020). Most researchers agree that NCLB is not meeting its intended effects, and research has found that since the implications of NCLB, many students, particularly in the nation's lowest-performing schools, have lauded exam-oriented learning, faced with teachers who are conducting test based lesson and giving importance only in core subjects such as science and math while considering other curriculum areas and activities (i.e., social studies, sciences, art, music, physical education etc.) as unimportant (Amrein & Berliner, 2002; Haney, 2000; Nichols & Berliner, 2007). Critics claimed there was a lack of evidence that teachers who left after the implementation of a new system would be the weakest teachers or would replace them with more effective teachers (Robertson-Kraft & Zhang, 2016). In addition, critics have warned that assessment systems that rely heavily on test score data can demoralize teachers and encourage effective teachers to leave the profession (Baker et al., 2010). Still, in the NCLB Act, students' background characteristics or the exact time they entered a particular school were ignored. This sentiment is laudable and hosts problems like making unfair comparisons. In addition, although providing effective instruction constitutes the primary goal of teacher evaluation, the scores of the students were insufficient in determining the effectiveness of teachers in teaching, and there was a need for a system that provides more accurate judgment about teacher competencies (U.S. Department of Education, 2007).

Although it was seen that the studies carried out to evaluate the teacher over the years had made a certain progress, many factors that still went wrong were frequently mentioned by the teachers working in the schools. To listen to this voice rising from schools' research were conducted and The Widget Effect report was presented. The failure of the evaluation systems to provide accurate and credible data about teachers' instruction performance is called as the "Widget Effect." It is used to describe the tendency of schools to assume classroom effectiveness is the same for all the teachers by denying the individual strengths or weaknesses (Weisberg et al., 2009). Based on a series of research conducted in 12 districts, four states (Arkansas, Colorado, Illinois, and Ohio) with 15,000 teachers and 1300 administrators, it was revealed the failure of the teacher evaluation systems in providing accurate and credible data for teachers'

performance (Weisberg et al., 2009). The results showed that school members, especially teachers and administrators, both recognize ineffective teaching in their schools, but these districts are having problems in identifying the poor or effective teachers. This problem occurred due to short and infrequent observations mainly conducted by principals without getting training. 62% of the teachers mentioned that they were not aware of the concerns of teacher evaluation before they were being evaluated. Most of the teachers were not receiving specific feedback after evaluations to improve themselves, that making them feel like they were being treated unjustly. Only 25% of the teachers reported participating in a single informal conversation with the principal about improving instructional performance over the last year. According to Weisberg et al. (2009), there is no single correct evaluation model, but we should produce credible systems by setting clear performance standards, including rating options used for describing the instructional performance, monitoring the judgments, providing frequent and regular feedback, linking the system with professional development, and supporting teachers who fall below the standards.

Within the growing dissatisfaction of NCLB in 2009 and results of Widget Effect, President Barack Obama announced an education initiative called "Race to The Top Program," in which it was offered states funding if they are willing to overtake their evaluation system, which is fair, rigorous, and transparent (Marzano & Toth, 2013). Throughout the Obama administration's this reform, a competitive grant program initiated an unprecedented wave of teacher evaluation reform across the country. In 2010 U.S. Department of Education proposed "A Blueprint for Reform" to state teachers will be evaluated to focus on recognizing, encouraging, and rewarding the excellence (Marzano & Toth, 2013). Through this blueprint in 2012 two significant changes were implemented through the Race to The Top Project, which are using measures of student's growth as indicators of teacher effectiveness and making rigorous measurements in the pedagogical skill of the teachers (Marzano & Toth, 2013). In this period, the development of the policy-based school and teacher accountability reforms included two critical transitions: first, the transformation of teacher observation systems from a personally reflective mentoring and capacity

building activity to a metric-driven assessment process; second, adding and highlighting the growth of student academic performance as a measure of teaching quality (Sloat et al., 2017). Like other reforms, this reform also encountered some different approaches during its implementation and due to the development of high-stakes, policy-based accountability reform, states started to use quantitative measurements by using standardized observational frameworks to evaluate teachers. Unfortunately, over time, these efforts have also largely failed to produce significant improvements in teaching and learning, according to the sources (Firestone & Donaldson, 2019; Lavigne & Good, 2019).

Although the implemented reforms led to considerable progress in the field of teacher evaluation and to take steps to improve by learning what is wrong, the process of evaluating with VAMs (value added models) could not be prevented during this period. VAM can be defined as high-stake measurements used to classify teachers' effectiveness according to the statistically measurable effects of teachers on their students' standardized test scores over time (Close et al., 2020). Every Student Achieves Act (ESSA) in 2015 is the US Congress's reauthorization of the Elementary and Secondary Education Act of 1965 (ESEA). The legalization of ESSA is a notable inflection point because it signaled the end of the era of high federal activity, which included two initiatives: Race to the Top and ESEA flexibility, which encourage the development and implementation of more objective teacher and principal assessment systems (Ross & Walsh, 2019). In 2015 (ESSA) signaled states can have more freedom in changing teacher evaluation policies and local control over teacher evaluation will increase (Klein, 2020). ESSA has also allowed states and territories to develop local teacher assessment systems that use alternative methods and measures to link and evaluate student development with teacher effectiveness (Close et al., 2020). In other words, states have more freedom through ESSA to determine the types of assessment and evaluation that can be used to determine teacher evaluations (McQueen, 2022). When this change was announced, it was unclear whether schools would continue to use the Value-Added Model's student achievement scores to predict teacher effectiveness. In fact, the study conducted by the National Council on Teacher

Quality (NCTQ) revealed that states did not make major changes after ESSA and continued to use VAM scores (Walsh et al., 2017). However, the new freedom that ESSA provides to states means they can move away from such high-stakes and appraisal models of accountability, especially those based on VAMs. Many states today offer more alternatives to teacher assessment by measuring the relationships between student achievement and teacher effectiveness, state teacher evaluation plans also include more formative teacher feedback, states also allow districts to develop and implement more specific teacher assessment systems (Close et al., 2020).

2.1.1.2 Pioneering Models Used in Teacher Evaluation

Changing policies and proposed reforms in the historical process have shaped teacher evaluation considerably. In this shaping process, many teachers' evaluation models have been developed by leading researchers. Of course, teacher evaluation models are still evolving in the light of recent studies but among these models, the most prominent ones have been tried to be explained by including them in the historical process.

In the 1990s, Charlotte Danielson and team members working at Educational Testing Service (ETS) developed a package of teacher-licensure examinations known as Praxis. Danielson worked on Praxis III and developed a system for the evaluators responsible for making judgments about the strengths and weaknesses of teachers (Toch & Rothman, 2008). Danielson noticed the teachers she trained as evaluators used the model to improve their teaching in those years. After that, ETS published her famous manual "Enhancing Professional Practice: A Framework for Teaching" in 1996. Through documenting empirical studies and theoretical research as promoting improvement in students, the learning framework was updated in 1997, 2007, 2011, and 2013. Throughout the changes in education and Common Core state standards, this framework was adopted in most of the states to envision active engagement of students in learning, deep conceptual understanding, thinking, and reasoning, and developing the skill of argumentation (Danielson, 2013). So that in 2013, to evaluate teaching for deep conceptual understanding, argumentation, logical reasoning, and making students take an active role in their learning specific additions included in this

framework. Today in this framework, there are four major categories, which are planning and preparation, the classroom environment (creating an environment of respect/rapport, establishing a culture for learning, managing classroom procedures, managing student behavior, organizing physical space), instruction (communicating with students, using questioning and discussion techniques, engaging students in learning, using assessment in instruction, demonstrating flexibility and responsiveness) and professional responsibilities. Other evaluation models which are also using Danielson's framework and rubrics, such as Teacher Advancement Program, Connecticut's Beginning Educator Support, and Training Program, The Cincinnati and Toledo, Ohio (school system evaluation models), and National Board for Professional Development are aimed to measure instruction for improving teaching (Toch & Rothman, 2008).

The System for Teacher and Student Advancement (TAP) system has been managed and supported by the National Institute for Excellence in Teaching (NIET) since 1999 (Culbertson, 2012, National Institute for Excellence in Teaching, 2022). This system measures teacher performance and provides ongoing, job-embedded, and collaborative professional development opportunities. System aims to improve teachers' skills through intensive support to master and mentor teachers by analyzing student data, creating achievement plans, leading professional development, and supporting teachers (Sloat et al., 2017). In addition, at the end of each year, teachers are also being scored according to their efforts to improve teaching, in other words, how they take responsibility for their teaching. TAP provides a comprehensive system for observing and providing feedback to honor and reward teachers for how well they teach. Ongoing training, mentoring and classroom support are also provided during school day as well as providing financial incentives for success (National Institute for Excellence in Teaching, 2022). The measurement and evaluation tools and the procedure of collecting evidence can be listed below.

- Each year, multiple classroom observations are conducted by trained and certificated evaluators (principals, administrators, master teachers, and mentor teachers).

- In-depth teacher conferences follow observations to examine strengths, weaknesses and prepare an improvement plan.
- In addition to teachers' classroom practice observations to provide evidence about student growth, a value-added score is also given to each teacher (Culbertson, 2012).

James Stronge's "Teacher Effectiveness Performance Evaluation System (TEPES)" has been implemented in more than 7625 schools and 20 countries since 2006. This system that aimed to support each teacher's continuous growth and development via meaningful feedback was developed for collecting and presenting data to document performance based on job expectations and guide instructional practice and the goal of the system (Stronge & Caine Tonneson, 2018). TEPES includes seven domains: professional knowledge, instructional planning, instructional delivery, assessment of/for learning, learning environment, professionalism, and student progress. Unlike other evaluation models, teachers are evaluated through formal classroom observations, documents like teacher artifacts, student surveys, and measures of student progress.

Measure of Effective Teaching (MET) project supported by Bill and Melinda Gates Foundation, conducted between 2009 and 2012. In this project, three thousand teachers in six urban districts contributed to the project to provide information. In the first report of this project in 2010, it was reported that student surveys are valued, and in the second report in 2012, it was mentioned that they value classroom observations more (Kane & Staiger, 2012). Consequently, student survey responses, ratings done by trained observers who gathered data from multiple classroom observation instruments and student achievement on state tests used as data sources to measure effective teaching (Kane et al., 2014). One of the critical points in the MET project is that observers are trained to enable them to evaluate the same competence areas in similar ways, and by this project, more than 900 observers were trained to score classroom videos using the various instruments (Kane & Staiger, 2012).

Classroom Assessment Scoring System (CLASS) was one of the protocols included in the MET Study developed to identify observable teacher-student interactions and focus on effective classroom instructions using observation tools (Rodriguez & Garza, 2014). Certificated observers use observation tools for fifteen minutes, and observation tools are designed to observe three domains: emotional support, classroom organization, and instructional support. These observation tools were used in more than 2000 classrooms. It was concluded that effective teacher-student interaction is crucial for the social and academic learning of the student and that professional development supports should be carefully designed and implemented to use this effective interaction at the maximum level (Office for Standards in Education,2018).

The in-service teachers' evaluation systems have also influenced the types of exams used to appoint the profession. The best example of this type of assessment is "Educative Teacher Performance Assessment (Detra)," in which the trainee can evaluate her practice-based teaching process. The first standard-based assessment EdTPA was developed in 2009 by the Stanford Center of Assessment, Learning, and Equity (SCALE). This assessment model is built on the National Board for Professional Teaching Standards model (CDE,2015). Today it is used nationally within 35 states and the districts of Colombia. EdTPA is a subject-specific performance-based teacher support system used for teacher candidates and requires teacher candidates to videotape their teaching and reflect on their planning, instruction, and assessment qualifications (De Voto et al., 2020).

The Center of Educational Leadership (CEL), working as a nonprofit service of University of Washingtons College of Education, developed a growth-oriented tool for improving instruction based on 5 dimensions of teaching and learning. This researched based tool named as “CEL 5D + Teacher Evaluation Rubric” and 5 dimensions were considered as “purpose”, “student engagement”, “curriculum and pedagogy”, “assessment for student learning” and “classroom environment and culture”. (Center for Educational Leadership, 2022). Rubric also includes “Professional collaboration and communication” which is based on activities and relations that teachers represent in and outside the classroom. The tool is used by the

schools and districts to create shared language for teaching and learning, scaffold the development of expertise and finally grow high quality instructional practices. Tool updated in years and 4.5 version released in 2020. In this final version promoting mastery-oriented learning, engaging students long term interests, valuing students' identities and giving feedback throughout learning process was emphasized (Washington Office of Superintendent of Public Instruction, 2022).

Being aware of the necessity and importance of teacher evaluation, these models and systems are implemented by different countries, states, and schools. However, the answer to the question of how fair teacher evaluation systems should be able to support and encourage quality teachers and quality teaching is still being researched, and therefore new models continue to be developed or existing models continue to be renewed. The pursuit of justice revealed the view that an important way to make the teacher evaluation system or models fair is to use student achievement scores to compare teacher achievement. This situation also brought the view that it would be fair to hold teachers more accountable to students, fee-paying parents, and taxpayers. (Idapalapati, 2019; Toch & Rothman, 2008). In fact, it is seen that the models that were very popular and applied for this purpose underestimate the student's taking responsibility for their own learning and, as a result, overestimate the responsibility of the teachers. Assuming that the effectiveness of teaching in a classroom is similar from teacher to teacher causes teachers to be understood as interchangeable parts, not individual professionals. However, it is not possible to talk about the effectiveness of teaching without determining the individual strengths and weaknesses of teachers (Idapalapati, 2019; Weisberg et al., 2009). While evaluating the teacher with only student success in the search for effective models is called as traditional methods, the tendency to develop new models to ensure the individual development of the teacher has started to gain momentum. It can be said that evaluation systems tend to benefit from many data sources such as systematic classroom observations, measurement of professionalism, measurements of school belonging, manager evaluations, test scores of each teacher's own class, to determine teachers' individual effectiveness (Glazerman et al., 2011; Ofsted, 2018).

2.1.2 Teacher Evaluation Processes in Other Countries

In “Teachers for the 21st Century Using Evaluation To improve Teaching,” the report revealed that OECD countries are embracing teacher evaluation increasingly, and 22 of the 28 countries surveyed stated that they included the issues related to teacher evaluation in their educational policies (OECD, 2013a). However, according to the report, some countries like Ireland, Portugal, Austria, Italy, and Spain have relatively poor evaluation structures, and teachers cannot benefit from any feedback given to improve their instruction. Furthermore, evaluation processes do not occur in policy frameworks in six countries (French Community of Belgium, Denmark, Finland, Iceland, Norway, and Spain). On the other hand, when these 28 countries are analyzed, it is seen that regular school-based evaluations were carried out in 17 of them, and 13 of them were made to understand that the experimental period of teaching was not completed.

In Finland, the Ministry of education has no role in teacher evaluation activities, and it is possible to talk about an evaluation focused on the professional development of teachers, which is group-based, reflective, and participatory rather than a systematic tool used for decision making (Tarhan et al., 2019). Teachers are evaluated depending on the content of the national core program and school development plans (OECD, 2013a). Finland's teacher evaluation process is opposite to the evaluation processes prepared for accountability because, in systems based on accountability, there are external evaluators in which the principal is employed. In contrast, the Finnish teacher evaluation process is carried out to cover a certain period in line with the development plan prepared by the teacher for herself regardless of the student, school, or the education system (OECD, 2009b). Similarly, the Swedish teacher evaluation process is not regulated by law, as in Finland. There is no formal procedure, and it is aimed to ensure the individual development of teachers individually by conducting self-assessment or speaking face-to-face with school leaders or colleagues through individual development dialogues (OECD, 2013a). This consultative and formative face to face conversation between teacher and principals or with colleagues teaching the same subject is a kind of discussion session conducted to evaluate the fulfillment

of the teachers' previous year objectives and conducted to establish further personal goals for the development of the teachers and school's needs (Williams & Engel, 2013).

Teacher evaluation is mandatory in three Asian countries, China, Japan, and Singapore, especially to excel in international PISA exams. The development of teacher evaluation framework in China and Japan is developed down to schools or districts; on the other hand, in Singapore framework is determined by the central education authority (CDE, 2015). In Singapore, teacher evaluation supported and encouraged teachers to realize the outcomes and behaviors specified by the Ministry of education. A competency-based teacher tracking system called "Enhanced Performance Management System (EPMS)" is used to improve teachers' performance by letting them reflect on their competencies and success and guiding them to plan and implement their professional development (OECD, 2009b). In such systems, teachers are assessed by a supervisor who could be the principal, vice-principal, or head of the department through an academic year, followed by regular meetings to review subject matter expertise, classroom management, and instructional skills. These systems support and encourage the teachers based on the data collected through observations, in-depth interviews, and focus group meetings (Steiner, 2010). Teachers' performance also is being assessed based on teachers' contributions to the school, how well they did the targeted duties, conducted cross-cultural activities projects. For identifying teachers' long-term potential and training needs, teachers are also evaluated for their "Currently Estimated Potential," which highlights areas that teachers will receive further training in a year (Williams & Engel, 2013).

In recent years, China moved away from the traditional understanding of teacher evaluation and has reformed teacher evaluation systems to improve the quality of teacher education. The ministry of education and schools are responsible for this teacher evaluation process, which is carried out to renew teachers' contracts and support their professional development (Taş, 2020). Principals and many evaluators evaluate teachers, and student test results were used as data for evaluating teachers (Liu & Zhao, 2013). Professional integrity, values, skills, competencies, teacher

diligence, and student success are used for evaluating teachers' performance. While doing so, teachers' self-evaluation, other teachers' evaluation reports, and parents' evaluations are used, and certificates that the teacher gains are also used as data source for evaluation (OECD, 2013a).

In Japan, in recent years, a system has been introduced for the renewal of the teaching certificate at certain intervals, and steps have been taken to improve education by teacher evaluation system which does not directly depend on students' scores in standardized exams and conducted for performance management and rewarding excellence (Kitamura et al., 2019). This system has been implemented in almost all schools and aimed to evaluate the teacher qualifications such as personal and professional values, instruction, organizing the classroom environment, pedagogical content knowledge and contributing to school development (OECD, 2013a; Taş, 2020). Evaluation results are also used for contract renewal supporting the teacher's career progression and increasing instructional effectiveness and overall job performance (Taş, 2020).

In Australia, Professional Standards for teachers was initiated by the federal government in 2009, and by these standards, professional knowledge, professional practice, and professional engagement were described in detail (CDE, 2015). Furthermore, these standards are also structured into four categories: graduate, proficient, highly accomplished, and lead to guide the teachers from diverse needs. The ministry of education and schools are responsible for teacher evaluation process, which is carried out to contract renewal, supporting the professional development of the teacher, supporting the teacher's career progression and increase instructional effectiveness and overall job performance (Taş, 2020).

2.1.3 The Development of Teacher Evaluation Process in Turkey

The evaluation of teachers' performance in Turkey extends from the Republic Period and all the educational services from May 1, 1920, attached to the Ministry of education. Although there was no supervision unit within the Ministry, three inspectors supervised teachers on behalf of the Ministry during this period. An

inspection unit was established under the Ministry of Education by establishing an inspection board on May 10, 1923. Turkey has carried out teacher evaluations as external audits by inspectors, and the inspectors started to be trained at the “Gazi Education Institute” in 1927 to conduct research and follow the developments in education in the world (Bilir, 1991). In 1938, it was aimed to guide the teachers by increasing the qualifications of the inspectors. Besides becoming an inspector, it was compulsory to graduate from Gazi Education Institute or graduate from the equivalent department abroad (Pehlivan et al., 2001). In 1969, the duties of the inspectors were briefly composed of auditing, professional assistance, and on-the-job training, review, and investigation (Buluç, 1997). With the continuation of external evaluations by the inspectors, the duties and responsibilities of the inspectors have been constantly updated over time. The boundaries of the fields of duty that inspectors were responsible for have started to be kept quite broad. The inspectors cannot be competent in every area, so they are responsible for various types. In many institutions to be evaluated, teachers' lack of clear competencies and responsibilities prevented successful evaluations. Therefore, it is impossible to mention that an effective inspector evaluation process was experienced until the 1900s.

In 1995, the Department of Education Research and Development conducted research called "teacher evaluation" in 117 schools to investigate the teacher qualifications as pedagogical competencies, professional competencies, personality traits, familiarity with educational technologies, and the capacity to make positive behavioral changes of the students. While the participants stated that each qualification area was equally important, it was concluded that eliminating the deficiencies by supporting the professional development of in-service teachers, who are lacking in these qualifications, is critical (Education Research and Development Department, 1995). Moreover, in 2001 the Department of Education Research and Development researched teachers, school principals, inspectors, provincial and district national education ministers. Research is conducted to determine who should assess the teachers working in primary schools, which time intervals, what kind of teacher

qualifications, and decide what to do with the performance evaluation results (Pehlivan et al., 2001).

Within the scope of efforts to improve performance competencies carried out by the Ministry of National Education (MoNE) and Education Development and Research Department in 2006, the professional development areas as "teacher competencies," "school administrators' competencies," and "school performance areas" were determined. In 2006 as a part of the "Performance Management Model at Schools" proposed by the Education Development and Research Department, it was aimed to evaluate teachers by all the stakeholders in the school with a system based on multiple data sources in the performance evaluation process. The impact of each stakeholder was determined as follows: School principal (50%), self-assessment (15%), students (1-5th grade student opinions 10%, 6-11th grade student opinions 15%), parents (1-5th-grade parents' views 15%, 6-11th grade parents' views 10%), head of department or teacher colleagues (10%) (Bozan & Ekinici, 2019). Considering the determined rates, the effects of school principals in the process are higher than other evaluators, which increased the importance of school principals in the evaluation process.

For many years, teacher evaluation continued to be carried out as an external audit with inspectors appointed by the Provincial Education Inspectors Department. However, in 2014, inspectors, under the guidance and supervision department of the Ministry of National Education, guided teachers, carried out inspections to improve the areas in which they were inadequate, and after the inspections carried out in these years, the inspectors, teachers, and school principals came together and prepared a development plan (Taş & Bıkmaz, 2019).

Duties, guidance, inspection, examination, and investigation roles of the inspectors and assistant inspectors are specified in the Regulation of the Directorate of Guidance and Inspection of the Ministry of National Education and the Directorate of Education Inspectors. (Ministry of National Education, 2014). In this system, a team of three to four investigators had undertaken a review visit for a maximum of three days. During the visit, they made classroom observations and interviewed school staff, parents, and

students on the school council. At the end of the inspection, inspectors met with the school administrators and teachers to share the inspection results. The report, which includes a list of suggestions that identify areas for school improvement, was shared with the school administration. The school management team is expected to develop a school development plan monitored by the provincial education inspectors based on the results obtained within a month (OECD, 2019). According to the regulation published in the official newspaper dated 24.05.2014, it has been prevented for the inspectors to observe the classes. In the official newspaper, dated 17.04.2015, it was explained that the responsibility of measuring the success, efficiency, and efforts of the teachers who have completed the candidacy process at the end of each academic year is given to the school principal working at the institution where the teacher is working (Taş & Bıkmaz, 2019). With these changes, inspectors only audited the school and did not conduct in-class audits. Within these changes' teachers' perception of fear towards inspectors has still not changed (Özkan & Çelikten, 2017).

According to Çiçek, Sağlam & Aydoğmuş (2015), to improve the quality of education in Turkey, there is a need to strengthen the internal audit instead of external audit, involve teachers in the teacher evaluation process, and guide to teachers through evaluations. Being aware of this need and recently exterior auditing, Turkey has also started to highlight internal audits. As a result, many areas of competence for the teaching profession and indicators were identified in 2017. Unlike previous 233 performance indicators and the heavy and complex previous competencies, the new competencies became more measurable with only 65 performance descriptors. In line with the OECD team's reviews, it can be said that previous competence areas and performance indicators guided new criteria for teacher development and evaluation, besides the new teacher qualifications are also clearly linked to national learning objectives (OECD, 2019).

The "Teacher Strategy Document" covering the years 2017-2023 of the Ministry of National Education aimed to ensure the employment of highly qualified and well-trained teachers who are most suitable for the teaching profession. This document also noted that it aimed to provide continuous personal and professional development

opportunities, propose a positive perception of the teaching profession, and strengthen its status. To achieve the objective of "ensuring the continuous personal and professional development of teachers," a periodic performance evaluation system to identify teachers' professional development needs was established. Furthermore, the quality of activities that target teachers' personal and professional development, starting from teachers' candidacy training, is planned to be increased (Directorate-General for Teacher Training and Development, 2017).

In February 2018, a draft of teacher performance evaluation regulations called "Performance Management System Module" was sent to external stakeholders. According to this draft regulation, teachers will be evaluated by multiple sources such as parents, students, principals, teachers, teachers from different and same branches, teachers themselves every year. Furthermore, all the teachers will take the Teacher Profession Qualification Exam every four years (MoNE, 2018; Özbek & Taneri, 2019). As a result of collecting the scores, the teachers would be gathered in four categories: A, B, C, and D. The teachers in the sub-category would participate in online and face-to-face in-service training. In addition, the appointment of teachers was planned by adding this score in addition to the Public Personnel Selection Examination (KPSS) score (MoNE, 2018). This module was activated on May 7, 2018, and when the module was opened and started to be piloted, it was criticized by many educators.

First of all, the purpose of this module, which was prepared to support the professional development of teachers by evaluating their performance, has been expressed incorrectly. It can be said that the primary purpose of teachers is to get a total score, to be classified in line with this score, and to examine the quality of the tools used. According to their performance outcome, separating teachers into specific categories (A, B, C, D) focuses on general weaknesses rather than individual professional needs by ignoring the multidimensional teaching process. On the other hand, the use of the evaluation results was not stated clearly, and evaluation should be used for particular formative assessment practices to support teachers (OECD, 2019).

Within this system, it was aimed to give a total score. However, when the system was implemented, it was revealed that %98 of the teachers had an evaluation score of 80 points or more because this evaluation has negative consequences beyond determining the professional needs (TEDMEM, 2018). Besides, General Proficiency Examination for Teaching Profession would be implemented and expected to be repeated at least every four years as part of the performance evaluation system, which will be insufficient to reveal the professional needs of the teachers.

In addition, an exam that will be implemented in this way will cause approximately one million teachers to become exam-oriented and create a new exam sector consistently. If the aim is to evaluate what teachers are doing in class and how they are carrying out the instruction, then behaviors that can be observed should be evaluated through determined performance indicators by observing work and actions. The most compelling evidence for performance indicators can be obtained by examining practices that directly observe the teaching and learning. Furthermore, by reviewing concrete documents that can be used as evidence for in-class observations such as lesson plans, teaching materials, activities, measurement materials (TEDMEM, 2018).

By this module, while evaluating the performance of a teacher, it is expected that school principals, a group of teachers from the same branch, teachers from other fields, the teacher himself, the parents, and students fill out the evaluation form. Parents and students may just have opinions about teachers; these opinions cannot be used as a judgment on the performance of the teacher and cannot be directly weighted by converting it to a score and used in total performance score (Danielson & McGreal, 2000; Isore, 2009)

In this performance assessment module, there are items such as "my teacher values us and shows love" or an item like "sensitive to the preservation of the natural environment and historical and cultural heritage." (MEBBİS, 2018). Of course, the teacher's value for the student and the sensitivity to the natural environment are significant. Still, it is not known whether the Evaluator will consider which actions or

practices are worthy or as evidence of sensitivity. Furthermore, such items are poor in measurability and can be interpreted differently from person to person. So, using such items when making judgments about the teacher's performance may lead to a misinterpretation of performance. On the other hand, performance indicators in the forms include value, attitude, and belief-based items and patterns like respecting, taking care, creating differences, etc. such as "A performance that contributes to the growth of students as individuals who are open to universal values, respecting national and spiritual values." This expression cannot be regarded as a performance indicator, or it cannot be measured by a questionnaire (TEDMEM,2018).

On July 20, 2018, the Ministry of National Education declared that this module and this system would not be implemented due to the inconveniences and criticisms. However, the effective assessment focuses on how well teachers support all students' learning and continually provides teachers with support and encouragement to improve their teaching competence. Therefore, summative development in many students' learning of the test results and the basis for receiving progress, giving teachers more reliable feedback is believed to be the critical creation of a fair teacher evaluation system (OECD, 2019). Furthermore, when used effectively, assessment can positively affect teachers' attitudes, motivations, and classroom practices, thereby helping to improve student's learning outcomes (OECD, 2013a).

When we look at the approaches followed in the historical process, the initiatives and the models developed in different countries, we can see that evaluating teachers is always an important issue, even though it has been made for different purposes. In addition, the historical process also gives us a clue that there is a trend towards more autonomous practices based on teacher development from country-wide assessment practices where student achievement is equivalent to teacher success (Murphy et al., 2013). The knowledge gained from the fundamentals and historical processes of teacher evaluation provides us with data on how we should not evaluate the teacher and forms the basis of how to make an effective evaluation.

2.2 Components and Characteristics of Effective Teacher Evaluation

Teaching is a profession because teachers are committed to behaving ethically and equally to all students, and it requires in-service training and continuity in post-graduate learning (Graham. & Berman, 2018). To change schools' currently functioning structure, it is vital to start a development process with the teachers and improve them professionally with well-developed evaluation results. As Stronge and Tucker (2003) mentioned, no educational reform effort can be successful without high qualified teachers, and as long as we have schools, the classroom teachers will be evaluated. This fact brings an inevitable solution: the need for high-quality evaluation systems to know if high-quality teachers exist.

Teacher evaluation is defined by researchers as an opportunity for teachers and administrators to collaborate and improve classroom performance, ultimately increasing student achievement (Reinhorn et al., 2017). In its simplest form, teacher evaluation can be thought of as the determination of a teacher's performance both in and out of the classroom by systematically collecting evidence and documenting the quality of teacher performance (Danielson, 2007; Stronge, 2006). The purpose of this systematic evidence collection process, how it is done and how the results are used shows the effectiveness of the evaluation. Besides, while preparing the right teacher evaluation models, it is crucial to distinguish teachers who are succeeding and struggling; furthermore, the suitable model for teacher evaluation should provide feedback and continuous improvement (Darling-Hammond, 2014). Most of the teachers who were considered successful in their profession had the opportunity to improve themselves by receiving feedback, working in collaboration with their colleagues, carrying out their own development professionally, and always focusing on the students. When the reputable resources related to teacher evaluation are examined, it was evident that the effective teacher evaluation process based on the development of the teacher includes formative evaluations with timely feedback, offers teachers the chance to participate in the evaluation process, provides clarity and consistency, collects evidence from more than one source, and considers contextual differences in the teaching environment (subject, grade level, class composition)

(Danielson, 2013; Darling-Hammond, 2012; Derrington & Brandon, 2019; Marzano, 2012; OECD, 2013a). In this section, the elements that should be included in an effective teacher evaluation process are given in detail under headings.

2.2.1 Role of the Teacher, Teacher Competency and Performance Standards

Teachers stay at the center of educational systems, and teachers' abilities and qualifications provide crucial contributions to students' learnings (Darling-Hammond, 2006). A study was conducted by Wright, Horn, and Sanders (1997) with 60,000 students from 3rd grade to 5th grade, and the results showed that the most affecting factor on student learning is the teacher. Teachers are reviewed as the most influential school-related force in student achievement so that teachers matter, and the work teachers do in the classroom also matters (Stronge, 2006). The success or failure of the teacher often goes beyond the classroom, and the teachers who do not reach the performance standards that they should achieve negatively affect the success of the students, the performance of other teachers, the reputation of the school, and the school administration (Jones, Jenkin & Lord, 2006). The fact that the difference in teachers also makes a difference in student learning makes it inevitable to invest in the professional development of teachers to train quality teachers (Darling-Hammond et al., 2017; Wiliam, 2018). Within the importance of the teacher, teaching has become a significant focus for policymakers and is considered a primary tool to be improved (Darling-Hammond, 2014; Stronge, 2006).

It is critical to distinguish teaching quality from teacher quality. Teacher quality is related to dispositions to behave in specific ways that refers to a collection of personal traits, understanding, knowledge, and skills of a teacher, while teaching quality is related to instruction that enables a wide range of students to learn (Darling-Hammond, 2012; CDE, 2015). Teaching quality can also be considered as it is in part a function of teacher quality. Evaluating teaching quality became more important late 20th and early 21st century when accountability took an essential place on the education agenda. Before implementing a teacher evaluation process, it is crucial to define conceptual consensus and shared understanding. With this significant rise, improper

use of competency or performance definitions has also increased considerably. It is essential to understand similarities between teacher competency and teacher performance, but it is also crucial to understand their differences (TEDMEM, 2018).

Competency is about having the knowledge, skills, and competence to perform a job or profession successfully and efficiently. According to the Ministry of Education Teacher Performance Evaluation Regulation Draft, teaching profession competency refers to the knowledge, skills, attitudes, and values designated by the Ministry that teachers should have to fulfill the teaching profession effectively and efficiently (MEB,2017). Competency alone does not show how the work is done or how efficiently and successfully it is done but just indicates that there is a certain level of capacity to perform a job (TEDMEM, 2018). On the other hand, performance relates to the process and results of an action, job, or task, and competency is a prerequisite for performance. Performance refers to the degree to which an activity or task is carried out effectively and is related to how effectively and efficiently these determined knowledge, skills, attitudes, and values are used in practice. Performance is a complex term that is a multidimensional construct, and it is essential to define performance to measure and manage the performance (Jones, Jenkin & Lord, 2006). The behaviors of the teachers transform the performance from abstract pieces to observable action (Brumback, 1988). Teacher performance can be regarded as teachers getting a task such as teaching, assessing, or monitoring; besides, performance means both the behavior and outcomes (Brumback, 1988; Jones, Jenkin & Lord, 2006). While defining the complexity of teaching, a list of tasks that identifies all different aspects of teaching and practices is considered as input; on the other hand, the results that teachers achieved in their work can be considered as output (Danielson & McGreal, 2000).

A teacher evaluation system should be evidence-based, and evidence should be based on clear and unambiguous criteria to define effective teaching (Danielson & McGreal, 2000). Teachers should be evaluated based on their ability to fulfill the teaching profession, which means setting clear performance standards in adopting comprehensive evaluation systems that are fair, accurate, and credible (Weisberg et

al., 2009). “How teachers perform in classrooms? How do they communicate with students? How do they communicate with students who have different cultural backgrounds? How do they cooperate with other teachers, administrators, parents?” are some of the questions that the answers could be found through setting behavioral criteria of qualification, which also be determined as standards. In many countries, standards term is used to state what teachers know and how they perform throughout effective teaching. Standards can be understood as checklist items of specific behavior, and this usage limits the meaning of teaching (Danielson & McGreal, 2000). Teaching standards should be observable, appropriate to provide feedback, and refer to teachers' competencies to make students learn. If the criteria are general and abstract, then the data which will support the teacher's development or accountability cannot be obtained (TEDMEM, 2018).

No matter what we choose to use as descriptors of effective teaching, it is crucial to consider that what is worth learning and how students should learn is changing. For this reason, it is necessary to evaluate the descriptors of effective teaching considering the changing process and the changing prospects. In other words, standards of effective teaching must reflect current best knowledge about learning (Danielson & McGreal, 2000). When determining the skills or standards that teachers should have, redundancy and complexity should be avoided so that the feedback given for different performance levels is meaningful (Danielson & McGreal, 2000). Another important point to consider when writing a standard is that the standards are expressed in performance terms such as what the teacher should do to support student learning and that they are concrete enough to guide teachers' observations in the classroom (CDE, 2015).

2.2.2 Purpose of Teacher Evaluation

Every stakeholder in charge of quality assurances, such as administrators, teachers, parents, inspectors, students, etc., plays different roles in ensuring improvement and accountability (OECD, 2009b). The purpose of the teacher evaluation needs to be defined clearly, and the more accurately the purpose is determined, the more the

evaluation system is built on solid foundations, and all stakeholders can correctly fulfill the role assigned to them while evaluating. Although the purposes of teacher evaluation may differ, the most common purposes are to improve teaching quality by strengthening teacher accountability which is summative in nature and improve teacher professional development, which is formative (CDE, 2015; Ford & Hewitt, 2020; OECD, 2009b; OECD, 2013a; Papay, 2012). Accountability is more concerned with the so-called causal effects of teachers on their students' learning (value added scores), as measured by the growth in students' scores on large-scale standardized tests over time, while assessments for professional development purposes are more concerned with assessment by observing teachers' practices (Amrein-Beardsley & Geiger, 2022). In addition, performance evaluation results are used as data for teachers about various decisions such as promotion, wage increase, performance-based payment, compulsory participation in training programs, and finally, to decide whether to continue to work or not (TEDMEM, 2018).

2.2.2.1 The Accountability Function of Evaluation Model

Accountability focuses on holding teachers accountable for their performance and aims to provide summative information about teachers' past practice and performance gathered through various sources (OECD, 2013a). Although there is no common definition for accountability, throughout the accountability process, individuals or institutions try to fulfill predetermined responsibilities or achieve goals, and they are obligated to provide an account of how they meet with the duties or goals (United Nations Educational, Scientific and Cultural Organization, 2017). A teachers primary professional responsibility seen as ensuring students learning thus, measures of students learning play a predominant role in teacher evaluation aimed for accountability. Strong and effective accountability systems which are much more than rating and evaluations are not solely based on standardized test scores provide valuable information to school and districts about performing levels of students that can be used by the school to guide teachers, give clear and comprehensive picture of the school to families (The New Teacher Project, 2016). For accountability systems to accomplish their predominant goal teachers should be accountable for helping students make

measurable progress against ambiguous learning standards (The New Teacher Project, 2010). Accountability took its place in the educational area with the No Child Left Behind act that promised all the students would achieve proficiency as measured by test scores which have still not been met over 12 years later (Kritt, 2018). Teacher evaluations for accountability purposes have been practiced since the 1990s. Holding the teacher accountable for students' results shifted attention to student achievement scores (Darling Hammond et al., 2012). Standards and accountability, defined as an educational reform, affect every aspect of schooling, teaching, and teacher education, especially in the United States. Accountability is carried out through the applied tests, and the data obtained from these tests has led to the production of books that are compatible with the tests and teachers to train based on test books, which has also affected the quality of teaching. On the other hand, holding students responsible for high achievements without holding the teacher accountable for their performance has revealed many other problems (Taubman, 2009). Providing incentives for teachers to perform better, entailing performance-based career advancements or salaries, bonus pay, or sanctions for low performance are the implementations of this type of evaluation (OECD, 2009b). When people know that they will be held accountable, they tend to take responsibility for their work, but it is not possible to talk about trust in the processes where there is punishment or reward, and accountability destroys trust when high stake external control is introduced by governments (Ehren & Baxter, 2021; UNESCO, 2017). The effectiveness of teacher evaluation for accountability depends on the correct setting of standards or criteria for teacher performance, the effective management of external and formal processes, and a cohesive and objective evaluation by well-trained and competent evaluators (OECD, 2009b). The accountability process can be trusted where stakeholders can work collaboratively, sustainable collaboration, and performance standards are established (Ehren & Baxter, 2021). Accountability is significant for improving education systems, but it should be a means to educational ends, not an end (UNESCO, 2017).

2.2.2.2 Improving Teacher's Professional Development

Most other professions build in a period of apprenticeship, like being an intern before being a doctor, but novice teachers receive little real assistance from another teacher in the school (Danielson & McGreal, 2000). One-shot sit-and-get workshops do not help teachers' professional development because the development of a teacher is not an isolated one-shot activity; instead, teachers need to take charge of their continuous professional development, and learning must be a part of the day-to-day routine for teachers (Attard, 2016). Therefore, it is essential to design evaluation systems to provide opportunities for professional learning for teachers, and teachers may improve their practice by engaging the activities required as a part of the evaluation process (Danielson & McGreal, 2000). Furthermore, to promote evaluation systems in which teachers are being evaluated fairly, accurately, and credibly the evaluation aim should be linked with professional development, and the core purpose of the evaluation should be to maximize the effectiveness of teachers, not documenting the poor performance evidence to dismissal (Weisberg et al., 2009). Stakeholders of the education field had common ground in the idea that teacher evaluation can be vital in increasing the focus on teaching quality and professional development of teachers since formative evaluation can also be used to raise teachers' self-efficacy, which is also a key component of being an effective teacher (OECD, 2013a).

Teaching requires a complex interaction between teacher, student, and content so that no one measurement tool can be used to make teacher evaluations. In evaluating teachers' performance process, the areas of professional development that needed improvement can be determined effectively from the data obtained from in-class observations, self-evaluation, peer review, teacher portfolio, parent, and student surveys (Almutairi & Shraid, 2021). Teacher evaluation for improvement purposes requires a non-threatening context, precise individual and collective objectives to improve teaching, evaluation instruments such as self-evaluation forms, classroom observations, structured interviews, and a culture of mutually providing and receiving feedback (OECD, 2009b). When the evaluation is conducted toward improving teacher practices, the evaluation process helps the teacher question and consider their

practices. Teachers are open to revealing their weaknesses because they expect to convey the necessary information to lead to more effective developmental needs and training (OECD, 2009b). This process may involve joint activities at the institutional level and individual learning, support, and mentoring activities. The aim of the performance evaluation is not that management systems apply to the workers nor forces workers to work much more than before to intimidate them. Good performance management can only be conducted within the collaboration of managers and workers. Collaboration among colleagues in the process of professional development and school-based small-group work is known to be more effective than traditional in-service programs. When teachers are entirely responsible and collaborate, there is little need for the administrators to enforce teacher evaluation (Kritt, 2018).

It is essential to provide a balance between the improvement and accountability functions of teacher evaluation. Still, because of the difference in their primary purposes, this balancing cannot be straightforward, and the tools and approaches used may vary (OECD, 2013a). For example, if the evaluation aims to improve teachers practice, then teachers will be willing to show their weaknesses, but on the other hand, if an evaluation model seeks to provide summative information to decide teachers' career paths, then teachers will try to hide their weaknesses (OECD, 2013a).

2.2.3 Evaluator

Teachers can be evaluated by themselves, by other teachers, or by principals. To answer the question of "How the Evaluator evaluates?" the qualifications of the evaluator should be well examined. It is significant for the effectiveness and reliability of the evaluation that the evaluators are competent and trained in the field of evaluation. The Evaluator's quality and ability should be monitored as the evaluation itself (Marzano et al., 2011). Schools and educational systems can change frequently, and evaluations need to keep pace with current educational policy and developments (Ofsted, 2018). Evaluation policies view teaching as complex and context-dependent, but there is a need for policies to construct evaluators who are well qualified and

distinguished by their ability in making professional judgments (Brandon et al., 2018; Cousins & Earl, 1995; Darling-Hammond, 2012; Darling-Hammond, et al., 2012).

To conduct an evaluation model that reflects variations in teacher effectiveness fairly and accurately, those running the evaluations, such as teachers, principals, assistant principals, etc., must receive rigorous training and support (Weisberg et al., 2009). The successful teacher evaluation system greatly depends on the in-depth training of the evaluators, and evaluators should have a range of characteristics and competencies (Cohen & Godhaber, 2015; OECD, 2009b; TEDMEM, 2018). According to the Ofsted report, it is crucial to maintain a high consistency between evaluators by providing training and refresher training (Ofsted, 2018). Calibration sessions are used to ensure that observers are sufficiently skilled to carry out lesson observations. Training should be intensive and ongoing, meaning that it should be repeated periodically (Weisberg et al., 2009).

In training, it is crucial to include the usage of prepared forms such as rubrics, non-structured or structured observation forms, etc. and videos taken from different teacher's classrooms to increase the confidence that the results can reliably indicate an evaluator's ability to apply the instrument as intended (Archer et al., 2016; Ofsted, 2018). Furthermore, in observer training, it is vital to include systematic approaches such as building observer capacity by providing intensive training programs, creating conducive conditions by removing the obstacles during the observation to promote a positive atmosphere (Marzano & Toth 2013). In addition, evaluators should receive training about some theoretical information such as a background in teaching; knowledge of educational evaluation theories and methodologies; knowledge about the concepts of teaching quality; familiarity with systems and procedures of academic and school quality assurance, including the role of teaching quality in school quality and the role of teaching quality in personal development; understanding of instrument development, including reliability and validity of observation and other assessment tools; awareness of the psychological aspects of evaluation; expertise with the quantitative rating of an assessment (OECD, 2009b). Furthermore, it is also essential to include some unique features of the specific evaluation model conducted in a

school, such as the aim of the model, performance standards to be evaluated, providing constructive feedback, designing, and delivering support teachers (Weisberg et al., 2009).

Principals take part in the profession with a range of educational and professional experiences, and recent educational policy developments, including teacher evaluation systems, raised the expectations for principals to improve school climate and instructional practices of teachers (Liebowitz & Porter, 2019). According to the research conducted by Liebowitz & Porter in 2019, there is a remarkable relationship between principal behaviors and teacher instructional practices. Being an instructional leader by supporting teachers' instructional practices by teacher evaluation, doing observation, and providing feedback to support the professional development of teachers also includes planning and developing education programs for teacher's professional development (Grissom et al., 2013). Therefore, it is essential to minimize the time on managerial and administrative tasks and maximize the time in favor of instructional leadership activities such as classroom observations and supporting the professional development of teachers (Liebowitz & Porter, 2019).

On the other hand, if only the principals were asked to do the observation in a classroom, in that case, the teacher would be passive, but if a school wants to enhance professional learning of teachers, teachers themselves should play a more significant role and be more active in the evaluation process. Evaluation systems such as 360-degree systems include the perspectives of other others (colleagues, parents, students, administrators, etc.) addition to teachers' perspectives and such systems are based on the idea that a teacher's skill may be seen from several perspectives, and it should also be examined from all different angles (Danielson & McGreal, 2000). Along with accepting the importance of the principals' involvement in the evaluation process, the fact that the principals do not have information about the observed teacher's field is also a critical situation (Brandon et al., 2018). For these reasons, teachers find it effective to be evaluated by peer teachers, which enables classroom insights and enhance evaluation feedback (OECD, 2013b). The effectiveness of peer review may vary in different cultures. However, peer review in school environments where

teachers see themselves as members of a professional profession can give beneficial results for evaluation. This kind of dialogue between teachers contributes to the development of teachers. When peer review applications are examined, this evaluation is not just about filling out a form; it seems to be a process involving exchanging views on observation, co-planning, work, and practices between teachers from the same branch or the field (TEDMEM, 2018).

While evaluating the teacher's performance, the self-evaluation process that the teacher carries out with the aim of empowering himself with his own participation becomes a necessity (Nikolic & Cabaj, 2003). Regarding the quality of the teachers, self-assessment can add much to the quality of teaching because teachers are perceptive of their strengths, weaknesses, skills, teaching, and the act of self-assessment includes reflections and growth. A result of self-assessment can be used by some supporting evaluations. Self-assessment provides encouragement for teachers to examine their practice (Marzano & Toth, 2013).

Parents have only limited means to know how teachers act following the expectations of students in the classroom, and they generally play an indirect role in the evaluation process due to their distance from the teaching profession qualifications, ignorance about what happens in the classroom, and they are valuing teachers' characteristics depart from student achievement (Isore, 2009). The roles of parents and students should not be in the "note giving" side of teachers' performance, but their opinions can be taken into consideration as data sources. Selecting parents and students as evaluators brings many problems in a process where the qualifications and competencies of the evaluators require such sensitivity (TEDMEM, 2018).

Using a well-designed student perception survey can provide reliable feedback about teaching instructional practices, which can be predicted by student learning. Surveys of parents and students can give much information, but students' age must be appropriate, and questions should be asked about class more than the teacher. Young children would not comment on the teachers' level of expertise, or which instructional strategies were suitable for the content. The data obtained from student and parent

surveys should be considered together with the data obtained from other sources. Converting the data obtained from parent and student surveys to points or numbers, then using these points to judge teachers is a problematic application in terms of accountability of the Evaluator. For this reason, in many countries, data gathered from parent and student surveys are just considered as "opinions (TEDMEM,2018).

2.2.4 Data Collection for Teacher Evaluation

As many data collection tools can be used in the teacher evaluation process, these tools are directly related to the purpose of the evaluation. For example, while the data collection tool based on an evaluation that supports teachers' professional development is classroom observation, the primary data source used in an evaluation model conducted for accountability may be the test score results of students. In OECD's report about appraisal systems of teacher evaluation in 29 different countries, it was outlined that the most frequently used evaluation instrument is classroom observation, interview/dialogue with the teacher, teacher self-appraisal, and portfolio (OECD, 2013a). In this section, the tools and data collection processes used in teacher evaluation models are included regardless of the purpose of the evaluation.

2.2.4.1 Classroom Observation

Teacher competencies are related to the knowledge, skills, attitudes, and values teachers need to possess, while teacher performance is related to how teachers can reflect these competencies to classroom practices. Classroom observations can be used very effectively to observe teachers' behaviors in implementing lesson plans, practices, and classroom management. Data that is obtained in-class observations can often provide concrete data on the direct performance of the teacher, which cannot be achieved with various evaluation forms or tests. When using evidence-based tools in observation, observers have more significant potential to improve classroom practice (Lynda et al., 2021).

Teacher observation is a straightforward way to provide feedback, and observations are done for two purposes: measurement and development (Marzano & Toth, 2013).

In addition, observation breaks through classroom walls and provides professional sharing, collaboration, joint implementation, and feedback (Cousins & Earl, 1995). Essential aspects of teaching occur when teachers interact with their students in classrooms, and classroom observations will provide the most accurate information about their performance. In addition, teachers can get intensive assistance that they do not have time to provide through observations done by several colleagues (Gordon & McGhee, 2019). Therefore, teacher appraisal is usually rooted in-classroom observation, and almost all countries use classroom observations to evaluate teacher performance (OECD, 2013a; TEDMEM, 2018).

Of course, there are some difficulties and deficiencies in making observations, but in the evaluation, systems created to help the development of the teacher, there are many classroom observations made during the year, and significant feedbacks are given immediately after the observation and evaluation (Darling-Hammond et al., 2012). The quality of the tools used while observing and the fact that the same teacher was observed many times by different people are the factors that directly affect the effectiveness of observation. Another critical factor is that the observer specializes in observing by receiving training, which many teachers think that competent observers should observe them (Bill & Melinda Gates Foundation, 2012; Brandon et al., 2018; Kane & Staiger, 2012; OECD, 2009b; TEDMEM, 2018).

Sometimes it will not be possible to observe every lesson that includes all types of behaviors. The observers need to see video recordings of the lessons, which are provided for the observers to use as evidence. Multiple observers can view the recordings and evaluate the teachers together (Marzano & Toth, 2013). On the other hand, sometimes short videos of the lesson could be used to provide evidence when enough data could not be gathered from the observations conducted previously (Marzano & Toth, 2013).

During announced observation, teachers behave differently than they usually do, which is a kind of error for observation. To minimize this error, using well-defined observation forms, including different competencies that teachers should have, is a

critical process. If the teachers behave differently for being observed, it would be difficult to feign competence for high-performance levels. Therefore, it is more beneficial to use announced observation or prerecord videos of the classroom instruction to observe an intended behavior. In addition, teachers can use different types of instruction. For instance, teachers may help the students interact with the new knowledge, deepen their understanding of new knowledge, or help students apply the knowledge in some other lessons. The chance of seeing all three types of behavior in unannounced observations is relatively small, so that one announced observation for each of three types of behavior would be used to teacher solve this problem (Marzano & Toth, 2013).

2.2.4.2 Individual Interviews and Debriefing with Observation

Interviews are used to gather descriptive data to develop insights on how interpretations can be a dominant strategy for data collection or how they can be used in conjunction with participant observation or other techniques (Bogdan, R., & Biklen, S., 2007). Most teacher evaluation models usually include individual interviews, which fosters the reflective discussions between the observer and teacher (OECD, 2013a). Structured interviews with teachers before and after classroom observations provides feedback to the teacher and provide information to the evaluator to understand the observation process and make meaningful judgments about the process. Supervision type also called "clinical supervision," includes pre-observation, observation, and post-observation conference steps. Being informed by data gathered from the classroom, and following a reflective dialogue makes this kind of supervision nonjudgmental and built on mutual trust (Gordon & McGhee, 2019). Classroom observations and interviews before and after the observation are essential steps that form the basis for the later stages of the system to determine the teacher's professional development needs at an individual level in a reflective way and guide the teacher in setting the performance goal (TEDMEM,2018).

Pre observation which can be conducted as a conversation is a reflective activity involving teacher collaboration for understanding their strengths and weaknesses and

helps to reveal pedagogical possibilities and solutions that teachers have not tried before. (Ahmad, 2020). Pre-conferences or pre-interviews provide the opportunity to communicate with the evaluation to make them understand that the observation process will be held is designed to provide support and encourage reflection and professional growth (NIET, 2012). During the pre-observation evaluator and the teacher being evaluated discuss the aspects of the lesson to be observed through the lesson plan to ask questions for the upcoming lesson (Gordon & McGhee, 2019). Pre-conference can also be used to understand the issues which may negatively impact the lesson (NIET, 2012). In pre-conference sessions, structured questions like "What prior knowledge should students have before this course/subject/concept? What are the teaching methods and techniques to be used during the lesson? What do you expect students to know at the end of the course?" can be used (TEDMEM,2018).

Debriefing can be conducted with conversation to enable the observer to explore their interpretations of what they observed (Zepeda & Lanoue, 2017). On the other hand, the primary purpose of post-conference is to allow the teachers to self-reflect on the lesson observed with the guidance of the observer by using leading questions (NIET, 2012). In the post-conference session, the evaluator and the teacher being evaluated discuss the strengths and areas that need development by passing over the criteria (Gordon & McGhee, 2019). The post-conference will provide a vital feedback opportunity for the teacher and sharing evaluators' observations with the teacher will allow them to monitor and rearrange the practices. In the pre-conference session, structured questions like "How do you evaluate the lesson in general? "How do you think the lesson went?" Did you reach the goals you planned during the lesson? If you will repeat the lesson, what have you changed? Why could this part of the course be well carried out/ do not carry out?" can be used (Lynda et al., 2021; TEDMEM, 2018).

2.2.4.3 Surveys/Questionnaires

Although it is used more frequently at the higher education level, it has been practiced for years to collect data from students who can observe the social, emotional, and instructional qualities their teachers bring to the classroom every day (Amrein-

Beardsley & Geiger, 2022). In some countries, student surveys are implemented to receive feedback on teaching practices and methods; in some other countries, these surveys are implemented at the institutional level (TEDMEM, 2018).

Student survey tools are cost-effective, can provide reliable data because they draw information from the people who have the closest and most in-depth knowledge of what teachers do in the classroom each day, namely students, and provide useful and actionable data such as observational feedback based on value-added approaches (Geiger & Amrein-Beardsley, 2019). On the other hand, students' prejudice against their teachers' race, gender, and other demographic characteristics, how students personally feel about their teachers regardless of their own instructional or pedagogical skills; expected grades of students; students' perceptions of course difficulty; the number of students surveyed; and students' willingness to take surveys seriously negatively impacts the data obtained from these surveys (Geiger & Amrein-Beardsley, 2019). Primary school students typically fail to assess what good teaching really means and looks like, and students at all levels may not be able to determine how well a teacher knows content (Amrein-Beardsley & Geiger, 2022). Furthermore, questionnaires in teacher assessment are often written in general terms and do not give much information about how teaching occurs since surveys cannot cover every behavior, they can only capture some aspects of teaching. In general, the results of such questionnaires show that although the school's success is extremely low, almost all teachers do great work (Weisberg et al., 2009). Questionnaires about parents' and students' satisfaction and opinions are seen as important data sources for evaluating performance but student and parent surveys would not be seen as entirely reliable sources for evidence, but they can be used to supplement other teacher performance indicators (Danielson & McGreal, 2000).

Given the strengths and limitations of student surveys, it is important to be careful when using survey data. The issues of reliability, validity and bias become all the more important if surveys are used to make decisions about salary, for example, but when student surveys are used for formative purposes, especially when teachers are allowed to think carefully about what the data does and does not show, how they should

interpret the results, and when encouraged, the feedback they offer can be a powerful driver of change (Amrein-Beardsley & Geiger, 2022).

2.2.4.4 Self-Assessment, Self-Reflection, and Self-Study

Self-assessment and self-reflection terms argue that teachers need to take responsibility for their evaluation, making decisions on the areas to develop and monitoring their learning (Attard, 2016). Reflect term means to think about something carefully and thoughtfully, and according to Dewey (1910), a reflection is an act of looking and describing one's experience with multiple ways of understanding. By reflecting on their work, teachers take a step back from their work with a point of view to identify and solve the problems or identify their success (Brandenburg et al., 2017). Through the self-evaluation, which follows reflections, teachers can make judgments, recognize their strengths, and identify weaknesses that allow them to formulate strategies for their improvement (Nikolic & Cabaj, 2003). Teachers can use data gathered from self-assessment to describe the changes in practice in a non-threatening and non-judgemental context to serve the learning needs of students (McCombs, 1997). Self-assessment encourages teachers to reflect on the factors that impact their teaching, such as personal, organizational, or institutional factors (OECD, 2013a). According to Marzano and Toth (2013), one of the simplest ways to provide precision and efficiency of gathering observation is starting with teachers' self-evaluation because self-reported information could promote teachers' reflections and show that their point of view is valued. Furthermore, self-evaluation enhances feelings of job security and helps teachers better comprehend the rationales behind classroom behaviors or activities (Nikolic & Cabaj, 2003).

A self-evaluation is an essential tool when the purpose of the teacher evaluation is based on the improvement of teacher practices, and usually, the improvement function of the evaluation is jeopardized, and the usefulness of self-evaluation instruments is being undermined (OECD, 2009b; Nikolic & Cabaj, 2003). Since it is generally thought that teachers cannot evaluate themselves objectively, this type of evaluation is avoided, but this also affects the active participation of the teacher in their evaluation

process. Other evaluation results can support self-evaluation reports because teachers have a better sense of their typical behavior. Furthermore, to detect bias within a teacher's self-evaluation, this evaluation needs to be compared to the evaluation scores recorded by other observers (Marzano and Toth, 2013). While identifying the quality of self-review, self-reviewed problems, outcomes of performance reviewed by principals and other teachers, scrutiny of pupil progress data can be used (Jones, Jenkin & Lord, 2006). Self-evaluation of teachers, collecting documents from their lesson plans, artifacts, student work, and then describing the teacher's performance throughout the evaluator's idea would put a teacher in a more active role (Danielson & McGreal, 2000).

2.2.4.5 Planning Documents, Teaching Artifacts, and Portfolios

A good planning document can be prepared with the great effort of the teacher, and its quality is related to the teacher's ability to conduct a successful instructional experience for students. The probability of utilizing effective classroom strategies and behavior is positively related to the efficiency of a teacher's planning abilities (Marzano & Toth, 2013). Planning requires thoughtful consideration of what are the students' interests, needs, backgrounds, and skills, understanding the subject area, engaging students with the content in a meaningful way, and content-specific pedagogy (Danielson & McGreal, 2000). Unit plans and lesson plans can be used to provide evidence for long and short-term planning. On the other hand, some teachers prepare excellent lesson plans which are brilliant on paper but carrying this plan out during the classroom teaching may not be as successful. Plans should not be used solely as evidence for effective teaching because planning requires complex skills, but planning may not mean that the teacher will be as successful in teaching (Danielson & McGreal, 2000; Marshall, 2013).

Artifacts represent an essential aspect of teacher performance and are as important as direct interactions for student learning (Danielson & McGreal, 2000). Artifacts such as worksheets, assignments, project directions, and materials created by teachers should be included in an evaluation system. Artifacts provide a window into classroom

life and provide evidence not only for classroom environment but also teachers thinking skills. Artifacts supported by classroom observations enable an evaluator to witness a teacher's plan to come to life for students. Teachers might provide classroom artifacts such as student response cards and the formative feedback given to students so that teachers could provide evidence for monitoring the student's development (Marzano & Toth, 2013).

Portfolios can provide strong evidence of classroom practices in performance appraisal, and they are not only an evaluation tool but also serve as reflective tools for teachers' classroom practices. Therefore, teacher portfolios are usually used as a complementary source for teacher evaluation, including pre-course preparations, lesson plans, teaching materials, a sample of student work, and commentaries on that work such as self-reflection sheets (OECD, 2013a). On the other hand, portfolios also include an action plan for instructional improvement, a self-assessment plan, and the teacher tracks to implement the action plan (Gordon & McGhee, 2019). Teacher portfolios include a vital data source for evaluating performance, and portfolios consist of all possible evidence to evaluate teachers, both summative and formative (Danielson & McGreal, 2000, Derrington, 2011). Furthermore, teachers' participation in conferences, training, and presentations can take part in their portfolios to provide evidence that shows how teachers put effort into their professional development and how they actively engage in the professional life by taking an active role (Danielson & McGreal, 2000, Derrington, 2011).

2.2.4.6 Student Achievement Scores-Value Added Measures

Learning is influenced by many factors such as students' skills, expectations, motivation, the structure of the curriculum, current and former teachers' influence on their learning, school climate, etc. Student learning outcomes, including school-based exam results and standardized test results, are an appealing measure to assess teachers' performance because the most crucial role of teaching is improving students learning. Thus, value-added models control a student's previous results and can potentially identify a teacher's contribution to a student's outcome scores (OECD, 2013a). Value-

added methods used for examining student learning gains into teacher evaluation are believed as they proved valuable data to examine teacher effectiveness of preparation programs, professional development programs, and various kinds of evaluation systems (Darling-Hammond, 2010).

Value-added measures or growth measures are used to estimate how much positive or negative effect teachers have on student learning by using statistical algorithms and standardized test results that integrated with other information about students. In value-added systems, predictive factors are used to determine how well the students achieve on standardized tests and this expected performance level of the student is used to evaluate the impact of each teacher (Weisberg et al., 2009). Predictions made with Value-added models are unbiased representations of a teacher's or school's contribution to student test success if the statistical model is correctly specified, statistical assumptions are met, and required measurement characteristics are valid, but researchers agree that these myriad conditions are never fully met (Everson, 2016).

Evaluating teachers' performance based on their student's achievement seems like a reasonable strategy but it is an important issue that needs to be clarified how to evaluate the teachers of the courses that we cannot measure by testing the students (Toch & Rothman, 2008). These models provide little formative feedback from the teacher and fail to demonstrate the teacher's influence on student's more significant developmental progress other than student achievement, which is tested annually (Master, 2014). For instance, most standardized tests used in teacher assessments today focus on low-level skills such as remembering or recalling information. Tests do not measure more high-level skills such as descriptive writing or the ability to think creatively or analytically, and they avoid art, music, and other subjects. In this case, the teacher who tries to teach students high-level skills falls into a disadvantageous position (Toch & Rothman, 2008). If we assume that student learning is a valid indicator of the quality of teaching, then it is expected to measure the learning validly like just not using multiple-choice tests but to measure the more complex form of learning such as written essays, conducted performance or experiments, designs, which should also be considered as the representatives of achievement (Danielson &

McGreal, 2000). The evidence of a student's achievement would not be limited by the score gathered from a test or exam, but students' writings, projects, presentation, or task which shows the skills they developed; student engagement in lessons should also be used as evidence (Danielson & McGreal, 2000). Student works can provide direct evidence of student achievement, and teachers should provide samples of student work selected to represent all ranges of ability and skill of the classroom.

Holding accountable teachers for the impact on achievement may not reflect reality because it is still doubtful that value-added systems can distinguish the contribution of teachers to students from the contribution of their parents and peers (Cohen & Godhaber, 2015). Furthermore, many factors influence students' learning out of schools, like socio-economic status, backgrounds, and socio-economic characteristics (Danielson & McGreal, 2000). In every school, some students are more challenging to learn than others, have low cognitive or behavioral features, have privileged backgrounds, and are bright or well prepared. This variable structure of the classroom environment results in many misjudgments of teachers' actual ability, and this unfairness in evolution could make teachers abandon challenging students (Toch & Rothman, 2008).

If student achievement would be used in teacher evaluation, the system should consider the baseline levels of student achievement and not the absolute level of performance but the amount they learned due to teacher effort (Danielson & McGreal, 2000). Nevertheless, there is still an essential consensus in the literature about two points. The first consensus is if performance results are mainly used for career decisions, student outcome should not be the only measure of teacher performance, and the second consensus is that using these results as an evaluation instrument for whole-school evaluation, not for evaluating teachers individually (OECD, 2013a). Value-added systems cannot be seen as the substitute for a comprehensive teacher evaluation system because they are less reliable in differentiating teachers in the middle range performance level, and value-added models can only be a helpful supplement to a credible teacher evaluation model (Weisberg et al., 2009). Knowledge about how well VA systems estimate teacher or school contributions to student test

scores is still evolving because there is much evidence that all VA systems' statistical assumptions are violated, and the question becomes one of "to what degree?" (Everson, 2016).

2.2.5 Use of Evaluation Results and Feedback Process

Once schools accurately and fairly evaluate the teachers, information gathered through evaluation should be used to modify teacher compensation systems, target professional development, recognize excellent teachers. People, including teachers, need to know how well they are doing, and performance evaluation process should include support by giving feedback and providing challenges together (Jones et al., 2006). The evaluation conducted for the improvement focuses on giving feedback to improve teacher practices through professional development that involves guiding teachers to reflect on and improve practices (OECD, 2009b). The professionally nurturing feedback for teachers seeking self-improvement should dig into evidence and include comprehensive, rich data that is systematically collected, prepared, and consumed (Amrein-Beardsley & Geiger, 2022). With such a feedback process, performance review systems become effective because positive reinforcement of an acceptable behavior increases the chances of the intended behavior being repeated (Jones, Jenkin & Lord, 2006).

The feedback given after the evaluation should be directed to the weaknesses or strengths of the teacher's lesson process so that development can be achieved; otherwise, the feedback given on the teacher's clothing, the layout of the classroom, or the answers of the students in line with a series of checklists does not make sense for the teacher (Culbertson, 2012). When the teachers were asked how they wanted to receive feedback, the teachers stated that they wanted to receive rapid feedback on students' learning in meaningful, non-threatening ways from people they trust (Guskey & Link, 2022). In addition, it is crucial to give teachers high-quality feedback based on accurate measurement of their instruction to improve teaching and learning (Kane et al., 2014). The use of valid tools, including specific criteria while giving constructive, detailed feedback, holding a reflective conversation with a sincere and

professional dialogue, and the fact that the feedback provided is entirely aimed at the development of the teacher strengthens the teacher's trust in the evaluation (Danielson & McGreal, 2000; Lynda et al., 2021).

Feedback should be given to the teacher immediately because if teachers do not immediately see evidence of a positive difference for their students, many will abandon a new practice and return to tried and trusted practices because continuing with an untested strategy is less likely for their students to learn well trust (Guskey & Link, 2022). Written feedback that is given after each observation ideally accompanied by a conference between teacher and the observer can enable observers to reflect consciously on their feedback and provide teachers with guidance that they can reconsider as needed (Putman et al., 2018).

Teachers are incredibly uncomfortable with receiving criticism about their behaviour performed in the classroom, but formative feedback aimed at the development of the teacher distracts the teacher from anxiety, eliminates the feeling of being judged, and frees the teacher (Gordon & McGhee, 2019). How this feedback is given is also very important in making this type of feedback useful. If teachers do not receive feedback in meaningful, helpful, and non-threatening ways, their practices will not change (Cherasaro et al., 2016). It is important to choose the positive things when starting the feedback, then to guide what needs to be improved and how to make the improvements, and finally to reflect the confidence that the person given the feedback will improve or change (Guskey, 2019).

Teachers need to open up their practice for review and constructive critique because that is what excellence requires (Reddy et al., 2016). To get feedback, teachers need to think about their teaching through reflective thinking, which means for teachers to think, analyze, and objectively judge their classroom activities to improve and develop teaching, and for accommodating change to their classroom implications (Liu & Zhang, 2014; Putman et al., 2018; Rushton & Suter, 2012). Reflective practices can be considered an essential part of lifelong learning; in other words, as a way of learning to teach, reflective thinking provides a flow of freshwater for the professional

development of teachers (Liu & Zhang, 2014). Furthermore, these practices are not preserved just for teachers but for support, development, and evaluation of teaching and learning (Rushton & Suter, 2012).

However, teachers rarely get as much useful feedback as they might want or not as much as they might need, to help themselves become accomplished teachers (Amrein-Beardsley & Geiger, 2022). Teachers are missing the opportunity to receive professional advice from their colleagues or supervisors, and they may be less likely than others to engage in focused professional learning and continuously improve their practice (Weisberg et al., 2009). Over half of the teachers in TALIS countries had never received any appraisal or feedback from an external source, such as an inspector. The internal appraisal was more frequent across countries, and 22% of teachers indicated that they had never received any appraisal or feedback from their principal, and 28.6% had never received feedback from other teachers or members of the school management team. Overall, 13.4% of teachers had never received any feedback or appraisal from any source (OECD, 2009a).

2.2.6 Validity and Reliability of the Evaluation

Validity is an essential concept in defining the quality of the evaluation, which means an assessment measure what is intended to measure and provides sound evidence in decision-making (Herman et al., 2011). Validation involves evaluating or justifying interpretations and includes uses of the scores (Herman et al., 2011). Therefore, rating scales used in the evaluation should be carefully constructed to minimize the most common validity problems like rater bias, halo effect, and leniency (Danielson & McGreal, 2000; Hollweck et al., 2019).

Reliability of the measurements is crucial because otherwise, data collected will paint an inaccurate portrait of teachers' practice (Kane & Staiger, 2012). For the reliability of the forms, the most problematic error is the occasion of the observation because classrooms are dynamic and complex settings, and the quality of student-teacher interactions can vary throughout one school day or over the school year. While estimating sampling error, it is important to do multiple observations and eradicate the

sampling error; the only way is to observe the teacher every day, which is not possible for the majority of the schools (Marzano & Toth, 2013). Therefore, many reliability problems occur due to the wrong type of sampling while observing teachers' instructional behaviors. In a lesson hour, teachers can guide students to interact with new knowledge, make practical activities to deepen the understanding or help students apply the knowledge to complex tasks. A teacher's level of using an instructional strategy may not be exhibited in one specific observation, or a particular strategy may not be observed during a single class hour. Many of the expected actions require a fair amount of time to implement. According to Marazano and Toth (2013), the probability of seeing each type of strategy in one classroom observation is extremely low. Observations should be made at least three different classroom sessions to observe enough to decide the teacher's behavior of all kinds and how the teacher uses strategies to eliminate the problem mentioned before. In an effective teacher evaluation model, it is needed to combine scores from multiple observers gathered from multiple observations to enable a high level of reliability. According to Bill & Melinda Gates Foundation Report (2012), a teacher's observation results varied from lesson to lesson and as well as observer to observer. The report also indicated that a focus on inter-rater reliability ignores the possibility that the teacher's practice may differ from one lesson to lesson or from one group of students to another. It is extremely important which data sources will be used in the evaluation based on multiple data sources and which field should be taken into account when evaluating the information obtained from these sources (Kahya & Hoşgörür, 2020). The key finding in teacher evaluation is a well-designed evaluation process used to measure teacher effectiveness that includes multiple measurements of teaching practice and student learning. An integrated teacher evaluation model that combines these measurements with productive feedback and professional development opportunities can increase teacher effectiveness and raise student achievement (CDE, 2015). Although teacher evaluation models have been under detailed examination for a long time, it is difficult to say that only one system is considered effective. But mainly including multiple measurements in measuring teacher performance, using multiple rating categories instead of binary ratings, creating a plan for the development of teacher performance based on these

ratings and its implementation is among the common features of effective systems (Steinberg and Donaldson, 2016; Steinberg and Kraft, 2017).

A reliable assessment means that giving a reliable indication of an observer's performance. Still, no matter the approaches used to ensure the reliability, it will not be accurate to say that the assessment is reliable just because different reviewers give the same score to the same behavior or response (Archer et al., 2016). Even a valid instrument is used in the evaluation, interrater reliability remains a concern and can be minimized by extensive training of the evaluators by developing a systematic approach for training (Hollweck et al., 2019; Marzano & Toth, 2013). In classroom observations, differences in raters' judgments may occur due to evaluation instruments, training of the observers, and the multifaceted nature of the observational (Wind & Jones, 2019). According to Marzano and Toth (2013), measurement error can occur due to the inaccuracy in identifying the type or the level of the strategies the teacher used in the classroom. That inaccuracy may occur by using a scale that is not specific enough to guide the observer or by the observer who is not well trained about the strategies used in the classroom. While providing consistency between raters, it is essential to include a clear conceptualization and description of the criteria or domains, and it is crucial to determine criteria clear enough for observers to look across the same set of indicators gathered by the instruments used in observation (Ofsted, 2018). To decrease the measurement error, multiple observers can do the ratings for the same lesson session by observing directly (live) or video recordings of the lesson and making concrete cut-points in the form for the observations (Marzano & Toth, 2013).

2.2.7 Credible and Trustworthy Environment

The success of systems depends on trust, and trust is a starting point as well as a result of doing the right things (Fullan & Hargreaves, 2016). The trust of those evaluated in an assessment system is a prerequisite for the correct and effective functioning of this system. Distrust may lie at the root of the tendency to evaluate everyone well and effectively in the evaluations made in schools. School climate and cultural norms are highly effective and have a decisive role in performance evaluation, and cultural norms

and expectations play a role, especially in colleague evaluations. Unfortunately, in teachers' evaluation of each other, there may be a tendency to evaluate each other "better" mutually or to evaluate them under the influence of school climate (TEDMEM, 2108). Evaluations could be credible if teachers and school administrators can capture what is going on all the time in a school, or at least they should be sure that teachers are doing the right thing (Marshall & Kim., 2013).

Undoubtedly, accessing multiple data from multiple sources is included in the process of collecting credible and reliable data in an evaluation system. Data regarding the teachers' classroom practices should be gathered from multiple sources collected over multiple points in time because the observation score can be prevailing with error for various reasons (Marzano & Toth, 2013). Therefore, evaluations should include multifaceted evidence of teacher practice using multiple evaluation tools such as self-assessment forms, classroom observations, and interviews (CDE, 2015). In addition, multiple sources, and tools such as observation, teacher's self-reflections, conversational learning time with peers, student achievement data, and feedback from other stakeholders can be used (Derrington, 2011). The Gates Foundation's MET project, which aimed to find out effective ways of teacher evaluation, started in 2009 and continued for three years (Bill & Melinda Gates Foundation, 2013). According to the project's findings, it was highlighted that effective teaching could be measured by using a mix of evidence sources and using a mixed evaluation measurement increases the understanding of the different components of effective teaching. In this project, three evaluation tools were investigated: classroom observation, student surveys, and progress in student scores (Bill & Melinda Gates Foundation, 2013). According to the Ofsted report, it is crucial to provide a whole-school approach by providing multiple observations and data sources to make reliable judgments about the teaching process (Ofsted, 2018).

If teachers believe in the need for change, modify their practices, work with supportive principals, and get instructional guidance, they successfully implement innovations to improve student learning (McCombs, 1997). An evaluation process that will not cause concern and fear and done through a culture of sharing will provide mutual

information flow (TEDMEM, 2018). Evaluation systems that encourage teachers to work with their colleagues and school leaders to identify goals and ways of assessing these goals will enable teachers to be a part of this system (OECD, 2013a). It is essential to listen to teachers and set classroom goals together to be fair between teachers. For example, if many struggling students are in one classroom, the goals should not be considered too high and not always met (OECD, 2013a). Furthermore, it is important to get opinions from teachers because they may not find it sufficient to see the results obtained in standardized test results as their own success, as their principals do. Although school leaders and principals prefer to use student success obtained from standard tests to predict teacher success, the development of students' attitudes, confidence in learning situations, self-efficacy and social-emotional learning skills can also reflect the success of the teacher (Guskey & Link, 2022).

2.3 Teachers Active Participation in Evaluation Process

For a change or activity to be accepted, successfully implemented, or sustained, everyone affected by that change or activity must be included and informed. Therefore, it is crucial to have all stakeholders' perspectives in the organizing process of teacher evaluation both for accomplishment and for the evaluation to produce accurate and valuable results (Shinkfield & Stufflebeam, 1997). The participation of a representative group of stakeholders in designing and developing personnel evaluation systems and determining evaluation criteria and standards increases the appropriateness and applicability of the system (Howard & Gullickson, 2009). Furthermore, encouraging the participation of stakeholders increases the common understanding and ownership. A performance evaluation system aimed at the professional development of teachers should be developed together with the teachers because teachers are not a tool of the school eco-system but an essential part of it. Considering that the most important stakeholder of a teacher evaluation system is the teacher herself, it is undeniable that their contribution to the development process of the evaluation system is also significant. To be more efficient, teachers can contribute evaluation systems features such as criteria, tools, the period they would believe the effectiveness of this system more (Danielson & McGreal, 2000). According to

Marzano and Toth (2013), increasing teachers' active participation in the evaluation process is one of the essential changes in the era of the effective teacher evaluation system.

Furthermore, as another part of the participation process, teachers need to make assessments of their professional performance by reflecting on their strengths and weaknesses in the lessons, rather than evaluating themselves with a rating ranging from "very little" to "very good." Thus, the teacher will have a say in their professional development and will be convinced that the system will be built on improving the performance, and the results will be used for their benefit (Attard, 2016; Brandenburg et al., 2017). Therefore, as part of the individual and professional development of the teacher, a reflective evaluation form, which includes strengths and weaknesses, and gives opportunity to express their opinion should be collected as additional data sources and should be considered together with the teacher's development plan (TEDMEM, 2018).

2.4 A School Specific Teacher Evaluation

In 2017 Ofsted hosted an international seminar with many experts who worked in teacher evaluation, and six teacher evaluation models were presented and examined in that seminar. A report was published after the seminar, and based on this report, experts agreed that it would be a mistake for schools to pick up an off-the-shelf model from elsewhere and apply it wholesale (Ofsted, 2018). A school that is unique and has its eco-system could need a specific evaluation. Schools should be considered with their teachers, administrators, curriculum, culture, etc., and teacher evaluation approaches that focus on professional development need to be designed in alignment with school contexts and objectives (OECD, 2013a). Teacher evaluation systems should be considered in the eco-system of teaching and learning (TEDMEM, 2018). The school usually conducts evaluations to identify individual teachers' strengths and weaknesses so that school leaders can make more informed choices about specific professional development activities for teachers at that school and in the context of that school's priorities (OECD, 2013a). As a starting point, it is essential to establish

an evaluation system based on the school autonomy, national curriculum, or culture of evaluation (OECD, 2009b).

The school management team usually carries out the teacher evaluation to improve teaching practices. The result of this evaluation system can be used to contribute the definition of school professional development plan used to improve the teaching process within the school (OECD, 2009b). Although all schools are part of the same system, each school's professional development needs and development plans will be different because of the different environmental conditions and the needs of the students and the teachers. When the school fails to support the teachers, difficulties can arise, and underperforming teachers may increase. A school can fail to support the teacher for some reasons such as being managed by weak or ineffectual leaders, systems for supporting teachers are inadequate, and the commitment of staff training and development is low (Jones, Jenkin & Lord, 2006). A list of teaching skills must reflect current knowledge of learning, and each school can expect different types of students as an outcome (Danielson & McGreal, 2000). Effective teaching can vary depending on the department or the school (Jones, Jenkin & Lord, 2006). Applying the best practices in some schools might not be transferred to another school which is different in terms of socioeconomic status, administrative differences, or teachers' effectiveness.

School-based teacher evaluation has the advantage of giving the school ownership of the evaluation processes and ensuring that the school carefully considers all aspects. It also provides that the school context (school's objectives and socio-educational background) is considered, and the use of internally devised instruments is more appropriate for teacher evaluation for improvement (OECD, 2009b). In this direction, at the beginning of every academic year, common goals related to the school's education should be determined and plans specific to the school should be established. Within these general targets, group-based and individual studies should be realized, and teachers should make their development plans following the school development plan (TEDMEM, 2018).

Focus schools are good examples to make this topic clear. States that receive federal waivers to the No Child Left Behind Act were required to implement designated "Focus Schools" reforms. Reforms in the Focus Schools, which were described as a high priority, contributed to the gaps in success. These Focus School reforms stressed the school's need assessments and provided technical assistance. Unfortunately, it was found that each of 3 years, Louisiana's Focus School reforms had no measurable impact on school performance. The improvement activities should be explicitly implemented in these schools, but the needs analysis by identifying the difficulties and support they need is neglected, and federal practices rather than local improvement are included (Dee & Dizon-Ross, 2019).

2.5 Research Studies on Teacher Evaluation

2.5.1 Research Studies on Teacher Evaluation and Participatory Approach

Growth-oriented teacher evaluation systems help teachers identify and reflect on their teaching to determine the highly developed practices and those that need further development by self-reflection, ongoing feedback sessions, and finally, a growth plan based on evaluations. A qualitative study was conducted with twenty-one teachers to analyze their first-year experiences in a growth-oriented teacher evaluation program in Franklin Regional School (Fowler, 2001). Findings indicated that teachers appreciated the self-directed and individualized nature of the system. Furthermore, through professional learning culture, including peer relations, relations with the evaluator, and professional learning activities, teachers experienced collegiality due to teamwork, reflective discussions on focused and sustained professional learning activities that enable teachers to apply theory to practice. On the other hand, teachers have suggested that the documentation and summative evaluation process be better explained and implemented more systematically by the evaluators; significantly, the feedback process should be revised to get meaningful comments. One crucial suggestion of the teachers was about stakeholder involvement meaning that teachers being involved in the design of the Franklin Regional Growth-Oriented Teacher Evaluation Plan to be clear about the elements of the program, to understand and apply

the rubric and how the aspects related with rubrics and summative evaluation. The teachers stated that not getting their views while planning and implementing this system created the feeling that they dictated something, and the management was challenging them.

To identify the most significant elements of teacher evaluation as perceived by K-6 teachers and administrators according to their lived experiences, DiGrazia (2018) conducted a qualitative study with eleven teachers and administrators from primary, middle, and upper grades. According to the study's findings, participants perceived evaluation for professional development, constructive criticism, peer collaboration, and informal visits for classroom observation as most pertinent to their evaluation. They also mentioned that by involving peers in the teacher evaluation process, providing teacher-specific and timely feedback about instruction was beneficial for their development. The teachers who participated in the research stated that they wanted to have more information about evaluation processes. They also mentioned that the administrators should involve teachers in developing these evaluation processes. The study concluded that the teacher would accept the process willingly if teachers had more to say in these processes in which they are evaluated, developed processes meet the needs of all parties and the teachers have more control over the evaluation processes.

Phenomenological research conducted with school leaders in Michigan aimed to uncover charter school leaders' authentic experiences in teacher evaluation processes and determine how to balance the accountability and professional learning components to support teacher performance (Evans, 2019). Research findings indicated that school leaders promise assessment systems that support the development of teachers and stated that teachers, because the autonomy granted allows schools to be evaluated following the context of the school and needs of the school, want more autonomy in teacher evaluation systems.

2.5.2 Research Studies on Problems in Teacher Evaluation

In the research conducted by Süzen (2007), with classroom teachers working in a private primary school, teachers' opinions about the performance evaluation system were obtained using the qualitative research method. According to the research results, most teachers stated that performance evaluations should be done for professional development purposes, and more than one person should do the evaluation. In addition, teachers noted that insufficient time, performing biased assessment, having problems in communication, and using evaluation results for purposes like comparing teachers, creating a competitive environment, holding responsible for student failure are problems experienced in the evaluation process.

In 2012 Bill and Melinda Gates Foundation conducted a study with the help of 3000 volunteer teachers to provide valid and reliable feedback to teachers for professional development and improvement. Bill and Melinda Gates Foundation analyzed how teachers and administrators could use teacher evaluations to improve teachers teaching skills (Archer et al., 2016). Two-thirds of American teachers had criticized the traditional evaluations methods, which assign ratings rather than changing practices. They believe that it did not reflect the complete picture of their classroom's practices. However, teachers stated that they could trust data gathered from fair and reliable measures. (Gates Foundation, 2018).

In 2016 Donahue, conducted a case study in a high school to evaluate Marzano Teacher Evaluation Model (MTEM) implemented in this school. MTEM aims to build teachers' pedagogical skills over time and increase expertise through feedback. Research findings revealed that teachers misunderstood this model as seeing it as an extension of their past practices. Nevertheless, they integrated the model to achieve enough ratings to continue their employment. Furthermore, findings also revealed that teachers described the barriers as a lack of trust in the intention of the observers. They also noted a need for more concrete examples to implement the strategies aligned to the model; they feel overwhelmed and feel like they are doing the instruction for the evaluator. They also added that they felt like they were getting scores rather than

guiding students to learn in observed lessons. Furthermore, findings indicated that focusing on scores rather than growth indicators created resistance for teachers to embrace the model willingly.

In a case study conducted by Fuller (2022), it was aimed to reveal how teachers perceived the effectiveness of the teacher evaluation process in a small school. The study conducted based on the assumption that if principles understand how teachers feel about the effectiveness of the assessment process, they will make policies and procedures applicable to address these perceptions. The study consisted of 20 teachers from a small, rural, central Illinois school. This research has shown that teachers are aware of the need for teacher evaluation but find the teacher evaluation process too formal and ineffective, with excessive paperwork.

2.5.3 Research Studies on Effective Features of Teacher Evaluation

Kimball (2001) conducted a case study to evaluate Danielson's Framework for Teaching Model implemented in two school districts with teachers and administrators from different school levels. Danielson's framework is being implemented based on standards and attempted to provide an evaluation framework that enables substantive feedback to enhance professional practice. Findings demonstrated that this framework changed the nature of teacher evaluation practices by using a set of teacher performance standards, allowing a more structured process, providing opportunities for teachers to get feedback and dialog, basing evaluation decisions on multiple sources of evidence. Other research findings revealed that teachers generally understood the standards, moderately accepted them, and perceived them valid and fair. Teachers and administrators agreed that this teacher evaluation system focused on teacher growth and accountability. Teachers find the system influencing by providing feedback to teachers about instruction, professional development, and teaching by dialoguing with the evaluators.

A case study was conducted to investigate, analyze, and describe K-6 teachers' and administrators' perceptions of the teacher evaluation process utilized in an elementary school and factors that impact teacher practice (La Masa, 2005). Findings revealed

that the teacher evaluation model would be more productive in improving teacher practice if the evaluations were done in a more collaborative and reflective environment, especially with peers. They also added that the evaluation would be more productive if more observations were conducted with constructive feedback, observations were carried out at regular intervals when the time was used efficiently, evaluators were trained about the formative assessment process, and if the teachers play an active role in the decision-making process of the professional development and evaluation processes.

Icel (2008) conducted a study with teachers working in charter schools to explore teachers' and administrators' insights into the current evaluation model implemented for three years based on Danielson's Framework. Findings revealed that participants indicated that open dialogue, meaningful feedback, and communication between teachers and principals are essential for professional development. Furthermore, well-prepared rubrics with clear standards can be used to build trust between teachers and principals.

To determine the effect of performance control and leadership characteristics of principals in public and private schools and propose teacher evaluation processes to improve teacher performance, Ilgaz (2011) conducted a qualitative study. Throughout the findings, teachers suggested multifaceted and process-oriented evaluations conducted based on specific criteria, by multiple sources, and with different methods should be included in expected evaluation systems. Furthermore, teachers also stated a need for fair evaluation, which aims to guide teachers' training.

Marzano (2012) surveyed 3000 educators. In this study, a simple five-value scale is given to the participants within a score of 1 indicated a belief that measurements are the sole purpose of teacher evaluation, and 1 also means that the development should not be considered the purpose. On the other hand, a score of 5 indicated that development is the sole purpose of the evaluation. And 5 also means that the measurement should not be considered as the purpose of the evaluation. A score of 3 indicates that the purpose is equally split between measurement and development, 2

as measurement and development are essential. Still, measurement is dominant; on the other hand, 4 also means measurement and development are crucial, but development is dominant. 76% of the respondents selected a score of 4, and 2 percent chose 5. Hence, most of the respondents stated that the teacher evaluation should be used for both measurement and development but should be considered the more important purpose. Throughout this research, many of the participants stated that a single observation made for one year is not enough and that different people should evaluate the different course processes of the teacher.

Bige (2014) researched to define classroom teachers' opinions about school principals' supervising process with 386 primary school teachers. Results revealed that teachers are expecting guidance, feedback, determining the needs of teachers for their professional development process, reward, and appreciation, and increasing the quality of education from the supervision process conducted in their school.

Nelson (2015) applied a survey to elementary teachers in Illinois working in thirteen different elementary schools to gain insight into the teacher evaluation experiences and learn their opinions of potential changes to the teacher evaluation system. According to the study's findings, classroom teachers considered trained and competent evaluators, clear and straightforward standards, and adhered to procedures as the most critical and desirable components of a teacher evaluation system. Teachers also indicated that an ideal evaluation process should support struggling teachers, a self-reflection component, and frequent and regular feedback.

A case study was conducted with high school teachers, students, principals, and parents in a private school to examine teachers' opinions about accountability (Türkoğlu, 2015). According to the findings, there are difficulties and uncertainties in preparing teacher evaluation criteria and evaluating teachers' performances. Teachers play a minimal role in decision-making processes at school and are seen as more practitioners. In this study, it is seen that teacher accountability is caused by intense expectations and an internal accountability system in the school. The evaluation of teachers is carried out based on observations in the classroom and outside the

classroom by the department coordinator, and teachers are also evaluated through the questionnaires by the school principal, vice principals, department coordinator, students, and parents. Teachers considered the personal observations made by the administrators as the most exciting part of evaluations, and the biggest concern of the teachers about the evaluation is whether the institution will work with them the following year or not. The questionnaire is used as a trump card, and the results of the questionnaires make teachers feel anxious about the process. The camera system is used to observe in-class and out-of-class activities and constantly monitor teachers' anxiety levels. Other criticisms made by teachers are lack of feedback after the evaluation process, and the teachers stated that they expect positive reinforcements as a result of the evaluation process. The teachers' statements concluded that they needed motivation in particular; they considered principals to use positive communication and expected a supportive approach from parents and administrators rather than commanding words. The teachers stated that some teachers were treated privileged in this school, the warnings given to them were not provided to these teachers, and they considered this situation unfair.

Interviews with 32 randomly selected teachers and two years survey from 12,000 teachers were conducted to measure the perception of teachers on the clarity, practicality, and cost of the new teacher evaluation model called Chicago's REACH Students. Chicago Students (REACH) as a new system to evaluate teacher performance can be considered a reform to teacher evaluation (Jiang et al., 2015). This system was developed to focus on improving instructional practices. In the REACH system, multiple classroom observations are included. After conducting observations, conferences with teachers were planned to give detailed feedback for all teachers and provide guidance for instructional practice. This evaluation model also includes measures of student growth. According to the findings, teachers are generally optimistic about this new system, especially the observation component. Teachers found the observation process to be more transparent and provide useful feedback. But they are worried about using the student growth in their evaluation. Another finding

showed that beginning teachers consistently were more positive than experienced teachers.

Winslow (2015) conducted a study to explore teacher and administrator perceptions of the effectiveness of the feedback in improving teachers' instructional practices. A mixed-method study was conducted in a school implementing Danielson's teacher evaluation framework. Findings revealed that the frequency of being observed by multiple researchers is considered as the best practice. In addition, teachers favor face-to-face feedback than written feedback from principals after observations because they wanted to get the feedback with the help of a positive dialogue they established with the principals and discuss this feedback with principals.

A study conducted with 1420 teachers employed in K-12 schools which have utilized a teacher evaluation system since 2000, and it was aimed to investigate tenured teachers' perceptions of the effect of teacher evaluation on their quality and other factors that contribute to their improvement of instructional practices overtime (Jaffurs, 2017). The study's findings indicated that most teachers viewed the local teacher evaluation system as a tool for building effectiveness over time. Still, less than half of all respondents believe that the system assists teachers in a formative way for professional development. On the other hand, most teachers embraced the post-conference as the most impactful part of the entire evaluation system. In addition, they found the self-reflection processes teacher evaluation system as beneficial to support professional development and build more pronounced teacher effectiveness over time. The research also revealed that experienced teachers' self-reflection tendencies were inversely related to their years of teaching experience and inexperienced teachers were more prone to self-reflection. Finally, most teachers mentioned that professional learning communities, peer coaching, and mentoring were effective professional development activities.

A qualitative case study conducted by Donahue and Vogel (2018) in the Rocky Mountain school district, which has sought to develop a system of supervision and evaluation to support teacher effectiveness for more than a decade, aimed to examine teachers' perceptions of how the practice of supervision and evaluation affects their

daily classroom teaching practices. It was assumed that the findings of this qualitative case study will reveal which elements of an evaluation system should be added to existing models by reviewing of teachers' perceptions of the school district. The data were obtained through interviews with 30 teachers working in different teaching positions. Feedback, quality of relationships, assessment rubric, modelling, personal integrity, and self-reflection were identified as mechanisms enabling teachers to benefit from the existing system. These themes revealed the complexity of the system where multiple mechanisms must work in coordination for teachers to realize the benefits in their daily teaching practices. According to the results of the research, effective evaluation systems should provide a clear indication of instructional strengths with feedback for continued practice, as well as identify areas of relative weakness in a teacher's teaching and provide specific recommendations for change. Teachers stated that people who spend enough time in their classrooms and have sufficient teaching experience can accurately and effectively determine what is going on in their classrooms and also stated that effective feedback should be given in a timely manner. Finally, participants saw their own integrity and self-reflection as perhaps the critical component in determining the usefulness of an evaluation system regarding classroom practice.

Daghe (2018) conducted multiple case study research to examine the teacher evaluation models implemented in these schools and analyze which of the model's characteristics teachers and administrators perceived to be most effective in promoting professional development and student achievement. In all models examined, evaluators were getting training, especially about the rubrics used in the evaluation model. Participants believed that the training about using the evaluation rubric gave them better credibility and understanding of what to look for during the observations. Furthermore, participants stated that the use of feedback allows the teachers to improve the quality of instruction.

In qualitative research carried out by conducting focus group interviews with teachers, it was aimed to understand high school teachers' perceptions about the evaluation components (conferences, classroom observations, student growth measures, teacher

reflection practices) in supporting teacher professional growth and student learning (Fulton, 2019). The study's findings revealed that teachers were aware of the importance of the evaluation components, and they perceived teacher reflection as the most influential component of improving teacher practices. Furthermore, teachers requested more frequent observations and opportunities to review goals and professional practices; they also wanted fidelity in the tools used for evaluation.

McQueen (2022) conducted a study to examine the teacher evaluation system in terms of teacher effectiveness in a small public school in Northern California. The purpose of this qualitative case study was to determine teacher perceptions of the current teacher evaluation system in a small Northern California public school district. According to the findings of the study teachers stated that the teacher evaluation system should use defined goals and criteria, be formal, and provide effective feedback. The participants also emphasized that the teacher evaluation system can increase student success by providing valid data, improving teacher skills, and providing evaluator and teacher cooperation. The opinions of the participants in this study that the evaluation system can provide professional development to the teacher through teacher training, focusing on specific areas and working in cooperation with their colleagues are also remarkable.

2.5.4 Research Studies on Developing Teacher Evaluation Models

Collins (1999) conducted a case study to explore the types of supervisory practices in a private school. The researcher aimed to reveal how these supervisory practices are perceived in terms of strengths and weaknesses, impacts of these supervisory practices on teaching and learning, teacher development, and school improvement processes, and make recommendations to improve the current supervision system. The research was conducted with two education board members, principals, teachers, and students working in a private high school. Findings of the study indicated that both Ministry of National Education inspections and school-based supervisions, including classroom-based performance evaluation, beyond classroom performance evaluation and summative evaluation of teacher performance. Findings revealed that the participants

had serious concerns about how the evaluation was done regarding the clarity of the purpose, the criteria, and instruments used for evaluation, classroom observation process, feedback reinforcement, reliability, the effectiveness of supervisors, and their relationship with teachers, students, and parents. Collins recommended a teacher evaluation model in which teachers, heads of the department, have a word to say in the decision-making process of the evaluation procedures; the purpose of the evaluation stated clearly, all the procedures of the model explained in a written document. She also suggested a list of criteria that can be used by the school staff regularly, data collection procedures including pre- and post-conference before and after the observations, and data collection tools.

Kaplan (2019) conducted a qualitative study to develop a supervision model supporting the continuous professional development of teachers. Qualitative data was gathered from 81 participants (school principals and teachers working in primary and secondary schools). Overall, the proposed model aimed to ensure continuous professional development for teachers. Hence, in the model-specific criteria used for evaluation, multiple evaluations such as self-assessment of teacher, peer coaching, classroom visits were adopted, and results were used to ensure the teacher's professional development.

2.6 Summary of Literature Review

The joint agreement shows that teachers play an essential role in students' learning more than other issues, and every aspect of school reform depends on highly skilled teachers for its success. The magnitude of the teacher's influence on student achievement has led to a continuous change in teacher evaluation systems. The current global education trends and the necessity of equipping students with the necessary skills and preparing them for life when they graduate have necessitated the continuous updating of education policies and reforms. The view that the teacher is the most important factor affecting the education system and the holistic success of the student has been around for many years, and it is a fact that this view will not be abandoned in the coming years. Throughout the teacher evaluation history many attempts that

take their roots from past educational reforms such as Race to the Top, Teacher Incentive Fund, the Measures of Effective Teaching Project, No Child Left Behind waivers, and Intensive Partnerships for Effective Teaching (Dee et al., 2021; Donaldson & Woulfin, 2018; Garrett & Steinberg, 2015). Although education policies are constantly changing, one thing that almost all of them consider important and do not hesitate to invest is the professional learning and development of the teacher. It is inevitable that this point is considered important because if you want a good return on investment for teachers and teaching, you must attract, select, and develop teachers with high human capital in terms of knowledge, skills, and abilities (Fullan & Hargreaves, 2016).

While the effectiveness of teachers is evaluated according to certain characteristics, today the methods and techniques used by the teacher in the teaching process continue to be evaluated in terms of their effectiveness in gaining the expected skills for students. Today education is shaping the workers of the economy and nation with best schools, schools with best teachers and leaders will own the future (Stronge, 2018). Teachers need to develop knowledge, skills and attitudes that are necessary to adapt to the contemporary world's complex and uncertain circumstances and therefore teachers need to make instructional plans and decisions to teach effectively based on the needs of the learners. All teachers and classrooms are unique and there is no way to guarantee teacher effectiveness however, this does not mean that teacher effectiveness cannot be improved. Achieving effectiveness is a continuous process based on teachers' reflections, adjustments, and growth. Undoubtedly, many studies and research about teacher effectiveness have significant consequences, and it is possible to increase the effectiveness by taking lessons from these results. Before starting to work on increasing the effectiveness of the teacher in an institution, it is very important to determine what kind of weaknesses the teachers have in which areas by conducting an effective, specific, and comprehensive evaluation model. According to Haeefe (1993), the dominant model of teacher evaluation is in trouble because evaluation criteria lack validity, evaluators are not trained, evaluators award lenient ratings to teachers who are weak.

Teacher evaluation can be thought of as the assessing a teacher's performance both in and out of the classroom by systematically collecting evidence and documenting the quality of teacher performance (Danielson, 2007; Stronge, 2006). Teacher evaluation can be seen as an opportunity for teachers and administrators to collaborate and improve classroom performance, ultimately increasing student achievement (Reinhorn et al., 2017). In order for teacher evaluation to create opportunities for teacher development and thus be effective in increasing student achievement, it must have an accurate purpose and include valid evaluations in order to provide effective feedback. (Darling-Hammond, 2014). When the reputable resources related to teacher evaluation are examined, it was evident that the effective teacher evaluation process based on the development of the teacher includes formative evaluations with timely feedback, offers teachers the chance to participate in the evaluation process, provides clarity and consistency, collects evidence from more than one source, and considers contextual differences in the teaching environment (subject, grade level, class composition) (Danielson, 2013; Darling-Hammond, 2012; Derrington & Brandon, 2019; Marzano, 2012; OECD, 2013a).

As in all evaluation systems teacher evaluation also should be based on evidence. For teacher evaluation practices, evidence should be based on clear and unambiguous criteria to define effective teaching (Danielson & McGreal, 2000). For comprehensive evaluation systems that are fair, accurate, and credible, teachers should be evaluated based on clear performance standards showing their ability to fulfill the teaching profession (Weisberg et al., 2009). To develop teachers or keep them accountable teaching standards should be observable, appropriate to provide feedback, and refer to teachers' qualifications to make students learn. Furthermore, standards of effective teaching must reflect current best knowledge about learning (Danielson & McGreal, 2000). When determining the skills or standards that teachers should have, redundancy and complexity should be avoided so that the feedback given for different performance levels is meaningful (Danielson & McGreal, 2000).

By defining criteria, we can determine what to measure. Accurately determining how to measure the criteria or standards we will measure makes evaluation systems valid

and reliable. While looking for the answer to the question of how we will do the evaluation, it is important to first determine the purpose for which we evaluate and to make this purpose the goal of the whole process. Otherwise, all our efforts to create evaluation systems will remain inconsistent and meaningless attempts. Improving teaching quality by strengthening teacher accountability which is summative in nature and improving teacher professional development, which is formative are two common purposes used for teacher evaluation (CDE, 2015; Ford & Hewitt, 2020; OECD, 2009b; OECD, 2013a; Papay, 2012).

Accountability took its place in the educational area with the No Child Left Behind act. It was promised all the students would achieve proficiency as measured by test scores, but this promise has still not been met over 12 years later (Kritt, 2018). The effectiveness of teacher evaluation for accountability depends on the correct setting of standards or criteria for teacher performance, the effective management of external and formal processes, and a cohesive and objective evaluation by well-trained and competent evaluators (OECD, 2009b). The accountability process can be trusted where stakeholders can work collaboratively, sustainable collaboration, and performance standards are established (Ehren & Baxter, 2021).

On the other hand, to promote evaluation systems in which teachers are being evaluated fairly, accurately, and credibly the evaluation aim should be linked with professional development, and the core purpose of the evaluation should be to maximize the effectiveness of teachers, not documenting the poor performance evidence to dismissal (Weisberg et al., 2009). Teacher evaluation for improvement purposes requires a non-threatening context, precise individual and collective objectives to improve teaching, evaluation instruments such as self-evaluation forms, classroom observations, structured interviews, and a culture of mutually providing and receiving feedback (OECD, 2009b). When the evaluation is conducted toward improving teacher practices, the evaluation process helps the teacher question and consider their practices. This kind of evaluation may involve joint activities at the institutional level and individual learning, support, and mentoring activities.

After determining the qualifications of the teachers to be evaluated, in other words the standards and the purpose of the evaluation, it is important to develop the data collection tools to be used and to structure the data collection process correctly. Various data collection tools can be used in the teacher evaluation process, these tools are related to the purpose of the evaluation. For example, while the data collection tool based on an evaluation that supports teachers' professional development is classroom observation, the primary data source used in an evaluation model conducted for accountability may be the test score results of students. Data that is obtained in-class observations can often provide concrete data on the direct performance of the teacher, which cannot be achieved with various evaluation forms or tests. Teacher observation is a straightforward way to provide feedback, and observations are done for two purposes: measurement and development (Marzano & Toth, 2013). In addition, observation breaks through classroom walls and provides professional sharing, collaboration, joint implementation, and feedback (Cousins & Earl, 1995). Structured interviews with teachers before and after classroom observations provides feedback to the teacher and provide information to the evaluator to understand the observation process and make meaningful judgments about the process. Classroom observations and interviews before and after the observation are essential steps that form the basis for the later stages of the system to determine the teacher's professional development needs at an individual level in a reflective way and guide the teacher in setting the performance goal (TEDMEM, 2018).

Although observation forms and post-observation interviews are commonly used tools, teacher self-evaluation is also a frequently used data collection tool. A self-evaluation is an essential tool when the purpose of the teacher evaluation is based on the improvement of teacher practices. Since it is generally thought that teachers cannot evaluate themselves objectively, this type of evaluation is avoided and the usefulness of self-evaluation instruments is being undermined (OECD, 2009b; Nikolic & Cabaj, 2003). On the other hand, self-evaluation also affects the active participation of the teacher in their evaluation process. Other evaluation results can support self-evaluation reports because teachers have a better sense of their typical behavior.

Furthermore, to detect bias within a teacher's self-evaluation, this evaluation needs to be compared to the evaluation scores recorded by other observers (Marzano and Toth, 2013). Unit plans and lesson plans can be used to provide evidence for long and short-term planning. Although some teachers prepare excellent lesson plans which are brilliant on paper but carrying this plan out during the classroom teaching may not be as successful. Plans should not be used solely as evidence for effective teaching because planning requires complex skills, but planning may not mean that the teacher will be as successful in teaching (Danielson & McGreal, 2000; Marshall, 2013).

Another important issue in teacher evaluation models, in which what, how and for what purpose will be measured is determined, is planning how to use the results in the most beneficial way. Once schools accurately and fairly evaluate the teachers, information gathered through evaluation should be used to modify teacher compensation systems, target professional development, recognize excellent teachers. The professionally nurturing feedback for teachers seeking self-improvement should dig into evidence and include comprehensive, rich data that is systematically collected, prepared, and consumed (Amrein-Beardsley & Geiger, 2022). It is crucial to give teachers high-quality feedback based on accurate measurement of their instruction to improve teaching and learning (Kane, et al., 2014). The use of valid tools, including specific criteria while giving constructive, detailed feedback, holding a reflective conversation with a sincere and professional dialogue, and the fact that the feedback provided is entirely aimed at the development of the teacher strengthens the teacher's trust in the evaluation (Danielson & McGreal, 2000; Lynda et al., 2021).

It is also especially important to ensure the validity and reliability of teacher evaluation systems that contain so many variables, which is closely related to the trust that teachers have or will have in this system. Undoubtedly, accessing multiple data from multiple sources is included in the process of collecting credible and reliable data in an evaluation system. Data regarding the teachers' classroom practices should be gathered from multiple sources collected over multiple points in time because the observation score can be prevailing with error for various reasons (Marzano & Toth, 2013). Therefore, evaluations should include multifaceted evidence of teacher practice

using multiple evaluation tools such as self-assessment forms, classroom observations, and interviews (CDE, 2015). Moreover, the fact that evaluators and evaluated persons are competent about teacher assessment through regular training also ensures that assessment models are fair and reliable. To conduct an evaluation model that reflects variations in teacher effectiveness fairly and accurately, those running the evaluations, such as teachers, principals, assistant principals, etc., must receive rigorous training and support (Weisberg et al., 2009). The successful teacher evaluation system greatly depends on the in-depth training of the evaluators, and evaluators should have a range of characteristics and competencies (Cohen & Godhaber, 2015; OECD, 2009b; TEDMEM, 2018). All in all fair and effective teacher evaluation models, which providing timely and meaningful feedback to teachers, including observations done throughout multiple sources and training educators to become expert evaluators in order to evaluate teachers effectively, are needed (Bill and Melinda Gates Foundation, 2010; Darling-Hammond et al., 2012; Darling-Hammond, 2014; Dee et al., 2021).

Teachers especially care about having a say in the planning and implementation of teacher evaluation models. What is known about how evaluation experience can change teacher effort and effectiveness from teachers' perspectives is relatively limited (Tuytens & Devos, 2013; Taylor & Tyler, 2012). Teacher evaluation systems will be more effective if the stakeholders participate in designing process of the evaluation (ESSA, 2019; McQueen, 2022; Paufler et al., 2020). Stakeholder involvement meaning that teachers being involved in the designing process of the teacher evaluation plan to be clear about the elements of the evaluation, to understand and apply the rubrics, to be clear on how the elements of the evaluation related with rubrics and summative evaluation. In recent years, collaborative and participatory teacher evaluation models continue to be developed to improve professional practices in schools (Darling-Hammond, 2014; Lillejord et al., 2018; Shulha et al., 2015).

CHAPTER 3

METHOD

This chapter contains information about the research design, research questions, context of the study, data sources, data collection instruments, data collection procedures, and data analysis procedures. In addition, methods used for the research's trustworthiness and the study's limitations are also given.

3.1 Research Design

Qualitative research is generally used when there is a need for detailed and complex understanding to explore the problem or issue. While exploring this problem with a detailed understanding, it is essential to address the meaning that individuals or a group of people attribute to the social problem (Marshall & Rossman, 2015). Therefore, it is crucial to consider the uniqueness of individuals by understanding the ways that people view, approach, and experience the problem or issue, and this type of research is also carried out not to ignore the effect of interaction between individuals on the problem studied (Creswell & Poth, 2018; Ravitch & Mittenfelner, 2016). In qualitative studies, it is essential to involve multiple perspectives, identify factors involved in the research and finally develop a complex picture of the problem or the issue (Creswell & Poth, 2018). While carrying out studies related to the problem, data are collected in the natural environment where the problem has occurred. The themes are reached by conducting inductive or deductive data analysis (Creswell, 2013; Marshall & Rossman, 2015).

An important reason for the qualitative nature of this study was to understand in detail the current practices related to teacher evaluation in the institution where the research

was conducted and to understand the strengths and weaknesses of these practices in detail. Furthermore, it was also aimed to get detailed suggestions from the participants for an effective teacher evaluation model designed to be implemented in their institutions. Another reason for selecting a qualitative study was the need to get the individual and unique opinions of the participants many times since a participatory teacher evaluation model was aimed to be developed within the scope of the study. In addition, the qualitative research method was preferred on the grounds that different participants, multiple perspectives, and the interaction between the participants were important to develop a participatory teacher evaluation model. In this study, the experiences of teachers, principals and experts on the current evaluation processes and their opinions and suggestions about the needed model with different perspectives were also examined in depth with qualitative methods.

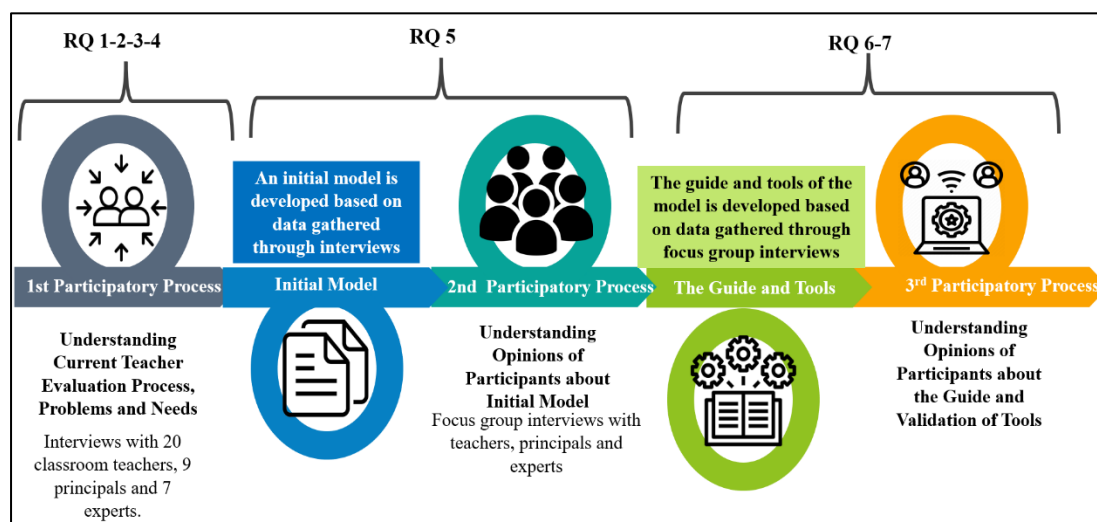
A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context and provides unique examples of real people in real situations by enabling readers to understand ideas more clearly than simply presenting them with abstract theories or principles (Cohen, Manion & Morrison, 2007). The case study is one of the qualitative approaches in which the researcher explores a real-life contemporary case or cases through a detailed, in-depth data collection process involving multiple sources of information (Creswell & Poth, 2018). A case study is appropriate when the problem or the situation cannot be defined apart from the context in which it occurs. In other words, case studies are preferred when the investigator has little control over the events and when the focus is on a continuing phenomenon with some real-life context (Yin, 2018). A researcher often selects a case that could be an individual, a small group, or an organization using purposeful selection. Thus, the primary concern is not generalization but developing an adequate description, interpretation, and explanation of this case (Creswell & Poth, 2018; Marshall & Rossman, 2015).

The teachers, principals, organizational cultures, the relationships between the employees, and the schools' dynamics are different. Therefore, for an effective teacher evaluation model that aims for professional development, it is crucial to be specific to

one school's unique ecosystem (Bülbül et al., 2013; Jones et al., 2006; OECD, 2013a; TEDMEM, 2018; Chen et al., 2021). Furthermore, primary school teachers play an essential role in shaping students' educational paths, so if teachers are to be encouraged, there will need to be a fundamental change in the educational climate (Webb, 2006). The case study method was chosen on the grounds that the teacher evaluation practices carried out in the school can only be understood in the context of the school itself. Furthermore, the case study is also chosen because the teacher evaluation model to be developed within the scope of the research should be unique for a school's own cultural structure, and the unique views of the participants should be reflected in this model. Following the research aims and examining the literature-supported explanations, it was decided that this research would be a case study. A private primary school was determined as the case, and the research was conducted on that school in the academic year 2018-2021. Classroom teachers, principals, and experts working in this school constituted the research participants. The data is collected through semi-structured and focus group interviews. The overall design of the study is presented in Figure 3.1.

Figure 3.1

The overall design of the study



According to Leedy and Omrod (2016), research begins with a problem, and literature related to the topic of the study is reviewed to formulate the problem. The problem addressed within the scope of this study was discovered by examining the research in the field of teacher evaluation and the sources containing different perspectives. After

identifying the problem as the lack of an effective teacher evaluation model implemented in the school, the research questions were developed. As shown in Figure 3.1, a general strategy to answer these questions, also called research design, was formed. The research design provides the overall structure of the researcher's procedures, like how the data were collected and analyzed. In line with the established research design, data collection instruments were developed, expert opinions were taken, and pilot studies were conducted. Based on the expert opinions and the findings of pilot studies, the researcher formed the final versions of the instruments. Next, the data collection processes were carried out. Since this study aimed to propose a participatory evaluation model, participant opinions were gathered multiple times. Their views were reflected in each step in the development process of the teacher evaluation model. Data were analyzed using content analysis. Furthermore, trustworthiness procedures were assured while collecting and analyzing the data.

3.2 Research Questions

The purpose of this research was to plan, organize and design a participatory teacher evaluation model to promote teachers' professional development. This research is composed of three parts. The first part of the research aimed to provide an in-depth analysis of perceptions of classroom teachers, principals, and experts on current teacher evaluation and professional development practices at the private school and to examine the suggestions of the participants in-depth. More specifically, the research questions were:

1. What kind of teacher evaluation practices are carried out for classroom teachers at this private school?
2. How are teacher evaluation practices perceived in terms of strengths and weaknesses by classroom teachers, principals, and experts at this private school?
3. What are the recommendations of this school's classroom teachers, principals, and experts to develop an effective teacher evaluation model?

4. What kind of professional development processes are carried out at this private school, and what is needed?

The second and third part of this research aimed to develop a participatory teacher evaluation model to promote teachers' professional development with data collection tools. More specifically, the research questions were:

5. What are the opinions of the participants regarding the initial teacher evaluation model?
6. What are the opinions of the participants regarding the guide of the model?
7. What is the validity evidence of teacher evaluation tools?

3.3 Context of The Study

Qualitative studies are context-dependent studies. The researcher should describe the contextual features and their influences in detail to make sense of how the events, actions, and opinions are shaped by the uniqueness of the context (Maxwell, 2013). This research was carried out in a private primary school in Ankara province. Ten classroom teachers were teaching in the first grade, 15 teachers were teaching in the second grade, 16 teachers were teaching in the third grade, and 20 teachers were teaching in the fourth grade. Almost all the teachers working in the school had at least three years of teaching experience. The number of newly graduated teachers working in the school was low. Teachers who are newly graduated do not teach during their first year of school, they observe other teachers. While the teachers are recruited, they are subjected to a written exam and an interview. In addition, an exemplary lesson process is observed in order to observe how teachers teach in the classroom during the recruitment process.

Admission to primary school took place in two different ways. Students attending the kindergarten of the same school both in the fifth and sixth years were entitled to study directly in the first grade of the primary school. No external students were accepted to study in the first year. Therefore, the number of 1st-grade teachers (n=10) is lower than in the other grades. On the other hand, students who have attended kindergarten and

primary school in another school were subjected to an exam. They were eligible to study from the second grade if they passed the exam. Participants believed that the students in this second group who earned to study at this school with an examination were more successful academically than those from the kindergarten. Some teachers reported that this situation affects the teacher's class average score and, hence the principals' view of the teacher's success.

A primary school principal, and twelve assistant principals are working in this school. Ten assistant principals working on each grade level separately. The school principal has classroom teaching experience for more than 25 years, and most of the assistant principals have classroom teaching experience. In addition, there was an assistant principal in each hall where the teachers' classroom was located. Five curriculum development experts and three measurement experts were working in this school. Experts were working in collaboration with teachers to design lesson plans, instructional practices, and assessment processes. Experts were also responsible for implementing projects at that school and for the organization of in-service training from time to time.

The professional development activities at the school are generally carried out within the scope of in-service training periods. Teachers are asked about what training they want to receive within the scope of in-service training, and training on subjects specified by the majority rather than individual requests of teachers was provided. It is not determined how the teachers could apply what they learned after the training. Teachers also notify the school principal when there are different trainings they want to attend, and the administration supports the teachers participation in this training. In general, they were of the opinion that the school provides less support than in the past on issues such as tuition fees, travel expenses, or meeting accommodation if the training is outside the city. There has been no systematic teacher evaluation process implemented at this school so far. However, the school principal tries to observe each teacher's performance non-systematically once a year.

3.4 Data Sources

Deciding where to conduct and whom to include in the research, called sampling, is an essential part of the study, and even in a single case study, it is necessary to explain why this case has been chosen rather than others (Maxwell, 2013). Although it is appropriate to use probability sampling or convenience sampling in research, there is a third category called purposeful selection or purposive sampling, meaning that particular settings, participants, or activities are being selected deliberately to answer the research questions (Maxwell, 2013; Patton, 2015). In the case studies, it is crucial to choose the most significant case which is accessible and illuminates the research question (Creswell & Poth, 2018; Yin, 2018). In purposeful sampling, it is essential to select individuals who have experienced the phenomenon being studied (Creswell, 2015).

3.4.1 Features of the Case

This study employed purposeful sampling, and the case was selected following the aims of the research, the research questions, and considering accessibility. The reason for choosing a private school rather than a public school was that public schools are subject to the decisions taken by the Ministry of National Education rather than the implementation of the individual choices of the schools in the process of teacher evaluation. There were many private schools in Ankara, and this school took a participatory attitude during the development of the evaluation model. In addition to the school's voluntary participation in the research, it was also crucial for the school to be accessible to the researcher and for the research to carry out a long-term and detailed research process. Since the subject of teacher evaluation is a subject that teachers approach with suspicion and uneasiness, it was especially important for the participants to trust the researcher in order to conduct the research more efficiently and effectively. The fact that the researcher worked in cooperation with teachers and principals for a long time as a curriculum development expert in this chosen school enabled the participants to answer the questions with confidence and in detail. Due to the mentioned reasons, the following features (see Table 3.1) are defined for this case.

Table 3.1*The Features of the Case*

<i>Criteria</i>	<i>Parameter</i>
Type of the school	Private School
Institutional characteristics	In need of a teacher evaluation model Being open to developing a teacher evaluation model Accessible to the researcher

3.4.2 Selection of the Participants

In qualitative studies, it is important to include multiple perspectives of participants to focus on the meaning that the participants hold about the problem or issue (Creswell & Poth, 2018). Classroom teachers, principals, and experts (curriculum development experts, measurement, and evaluation experts) working at that private school were selected as the participants. The first data collection was conducted with 20 classroom teachers, nine principals, and seven experts to understand the current teacher evaluation process, problems, and strengths and get recommendations about the evaluation model that will be developed as the product of this research. While 20 teachers were selected, teachers from each grade level with different years of experience were asked to participate in the study (see Table 3.2). On the other hand, those who wanted to participate voluntarily from the principals and experts were determined. As a result, nine principals and seven experts at the school agreed to participate in the study.

Table 3.2*Teachers' Grade Levels and Years of Experience*

<i>Grade Level</i>	<i>Teachers</i>	<i>Year of Teaching Experience in the School</i>
1 st Grade	T2, T4, T10	0-5
	T3	10-15
	T11	15-20
2 nd Grade	T6, T12, T16	5-10
	T8, T14	10-15
3 rd Grade	T1, T9, T17	0-5
	T5, T7	15-20
4 th Grade	T19	0-5
	T13, T20	10-15
	T15, T18	15-20

The second data collection was conducted through focus group interviews to understand participants' opinions about the proposed model, which was prepared according to the first interview results. Focus group interviews were conducted to elicit participants' feelings, attitudes, or perceptions about the topic. In focus group interviews, participants have the chance to express their views based on their experiences (Puchta & Potter, 2004). While selecting the participants, it is essential to choose the ones affected by the situation, who have enough experience to contribute by expressing their opinions in detail, and who provide diversity to understand the issue from multiple dimensions (Bader & Rossi, 2002; Krueger & Casey, 2015). Furthermore, a focus group participant numbers should be no longer than ten, and a group including six to eight participants would be enough to allow sharing insights and providing all the perceptions (Hennink, 2014; Leedy & Ormrod, 2016). It is crucial to conduct different focus groups to get various perspectives and increase confidence in the merging patterns (Patton, 2015). In the focus group discussions of this study, teachers, principals, and experts came together in each focus group meeting. Furthermore, both the teachers who participated in the first interview sessions and the teachers who had not been interviewed came together, shared their opinions about the initial model, and made suggestions about what could be added. Involving participants

who had different experiences in developing the proposed teacher evaluation model and participants with various duties and responsibilities that may be affected by the use of this model increased the interaction in the focus group meetings and enabled the researcher to get more in-depth opinions. In order for the discussions to be effective, the number of participants in the group was limited to seven. Interviews were conducted with four different groups (see Table 3.3).

Table 3.3

Focus Group Participants

<i>Focus Group</i>	<i>Teachers</i>		<i>Principals</i>	<i>Experts</i>
	<i>Participated in First Data Collection Process</i>	<i>Did Not Participate in First Data Collection Process</i>		
1st Group	T9 T13	T21 T22 T23	P1	E3
2 nd Group	T8 T11	T24 T25 T26	P3	E1
3 rd Group	T5 T10	T27 T28 T29	P2	E6
4 th Group	T3 T15	T30 T31 T32	P5	E7

The third data collection process was conducted to get participants' opinions about the final version of the model and get their views for validating the tools prepared for the data collection of the evaluation model. The guide, including the model's explanations and the data collection tools, was sent to all the participants via email to get their opinions.

3.5 Data Collection Instruments

In qualitative research, data are collected by tools designed by the researcher to examine documents, observe behavior, or interview participants (Creswell & Poth,

2018). So that the researcher developed both the interview and focus-group interview forms.

3.5.1 Semi-Structured Interview Forms

Interviews are commonly used in case studies to gather in-depth explanations by asking how and why questions to the participants (Yin, 2018). Interviews are conducted to find out things we cannot directly observe, like feelings, thoughts, intentions, perspectives of participants, and meanings attached to problems or issues (Maxwell, 2013; Patton, 2015). Interview schedules generally allow for open-ended responses which are flexible enough to collect data on unexpected situations and gather descriptive data in the participants' own words to develop insight into the participants' interpretations (Bogdan & Biklen, 2007). According to Mason (2002), qualitative interviewing should be used if the researcher suggests that the participants' perceptions, knowledge, views, understandings, interpretations, experiences, and interactions are meaningful sources to explore research questions. In this research, to answer research questions, gain a deep understanding, and develop insights into the participations, interview schedules were developed to collect data from teachers, principals, and experts on general and specific teacher evaluation issues. The process of developing the interview form took place in three stages. These are "writing the questions," "asking experts' opinions," and "pilot study."

In the question-writing process of the interview schedule, it is essential to consider how the situated knowledge could be generated (Mason, 2002). Focusing on lived experiences of the participants enables the researcher to generate the situated knowledge in an appropriate context (Mason, 2002). The interview schedule of this study includes questions asked based on lived experiences of the participants. While writing the questions, the researcher took the research aim and research questions as the basis. Interview forms developed for teachers, principals, and experts are similar in terms of the scope of the questions.

Probes can be called reminders for the researcher to ask for more information or ask for an explanation of the ideas (Creswell, 2015). In the interview form, the researcher

prepared and used probes when more information or clarification was needed to answer some questions. For instance, the qualifications that were thought to be evaluated were explicitly asked to answer the question in case the teachers do not understand what is meant by the term “qualification” or to eliminate the possibility of not coming to their minds.

An expert in qualitative research, two researchers experienced in qualitative research, two curriculum development experts, and two measurement and evaluation experts reviewed the interview schedule. The researcher consulted for their feedback on the content and face validity of the interview schedules. Based on their feedback, the researcher added new questions and removed and/or revised some of the questions. Questions were reexamined and changed to be more precise and more understandable, and the researcher also altered some of the questions to create a more integrated structure. For instance, to receive an answer associated with teachers’ experience, the question “Let us talk about the teacher evaluation process used in our school. Can you tell me how this process works?” is transformed to “Can you give examples of the evaluation processes at your school? Can you share your experiences about the evaluation processes conducted?”

On the other hand, the question “How is professional development linked to teacher evaluation?” was reported to be unclear. Therefore, instead of this question, the researcher asked questions for the participants to explain what was intended by professional development, give examples of the professional development activities carried out at school, and associate this with teacher evaluation. Furthermore, the question “What kind of steps should be included in a teacher evaluation model?” was not understandable by the participants and will not provide meaningful data for research. Therefore, it was decided not to use this question in the schedule. After reorganizing questions, overall errors related to the language were corrected and made ready for the pilot implementation process.

It is important to engage in a pilot interview to improve skills in interviewing. A pilot interview helps adjust the interview schedule if some of the questions are not clear

enough if the interviewee chooses to answer shortly, that does not provide sufficient explanation or if the probes are not working well (Creswell, 2015). Before the administration of the interviews, the questions were piloted with two classroom teachers, one expert, and one assistant principal to see whether the questions were understandable and clear or whether the questions were working or not. While teachers were interviewed in teacher rooms and classrooms, principals and experts were interviewed in their own offices. All the interview sessions were audio-recorded with relevant permissions. After the piloting process, the researcher changed the structure and the order of the questions. The first interview form consisted of questions in five parts: (a) the demographic information of the participants; (b) the current teacher evaluation processes applied at this school; (c) the positive and negative sides of these processes carried out at the school; (d) the suggestions about the model to be developed; (e) the professional development studies carried out at the school and the needs. When the researcher asked the questions in this order, it was noticed that the participants tended to answer other questions from the first question, and they could not answer each question in detail. For this reason, it was decided to change the order of the questions. For example, in the first interview form, questions asked to understand current implications about evaluators, schedule of the evaluations, data collection tools, data sources, teacher qualifications were included in one section, while the positive and negative practices carried out in the school were included in another section. The participant was talking about both positive and negative features while explaining the current situation. In the final version of the interview schedule, the positive and negative aspects of evaluating the teacher qualifications are asked immediately after the current situation related to teacher qualifications. In the interview schedule, the researcher noticed that the answers given to the “For what purpose are the evaluation results used?” question were not detailed because participants did not understand the question. So that the researcher decided to add probes like “How are the results used for determining your strengths or need to be improved?” These changes were also reflected in the principal and expert interview forms. The final version of interview schedules for teachers, principals, and experts are presented in Appendix A, Appendix B, and Appendix C, respectively. The

"Descriptive Information" section of the final versions included five semi-structured questions for teachers, six semi-structured for principals and six semi-structured for experts. In the "interview questions" section of the latest versions, there are 33 semi-structured questions for teachers, principals, and experts.

3.5.2 Focus Group Interview Form

The purpose of conducting a focus group is to understand better how people feel or think about an issue, idea, product, or service (Krueger & Casey, 2015). Focus group interview forms are helpful tools for developing or evaluating a program, clarifying, or validating the results of the interviews, disseminating preliminary findings, or gathering feedback to refine the framework (Creswell & Poth, 2018; Marshall & Rossman, 2015; Patton, 2015). In this study, focus group interviews were conducted with four different groups to get the participants' opinions regarding the teacher qualifications to be evaluated, schedule, process, use of results, and the feasibility of the initial teacher evaluation model. The development process of the focus group interview schedule is explained in three steps. These are "writing the questions," "asking experts' opinions," and "pilot study." The schedule was designed to include open-ended questions as the aim was to review each dimension of the initial model by the participants and get their opinions about some unspecified points that arose during the interview sessions. In the introduction part of the schedule, the researcher made an opening speech, introduced the model, and introduced the participants in each group. In another part of the schedule, the researcher asked open-ended questions, including all the model dimensions. Furthermore, to clarify some unspecified points raised in face-to-face interview, questions like "You mentioned that the observers could not be objective, so to solve this problem, can more than one person enter the observation at the same time? How?" "When the interviews were examined, it was stated that the evaluation of the lesson plans of most teachers should not take place as they do not reflect the process. What do you think about using the lesson plan within such a model?" were also asked to the participants.

After writing the questions, two professors experienced qualitative research, two curriculum development experts, and two measurement and evaluation experts reviewed the interview schedule. The researcher consulted for their feedback on the content and face validity of the focus group interview schedules. Based on their feedback, the researcher changed some questions entirely and added more questions. For instance, “What kind of weighting do you suggest being assigned for each qualification? Which qualification is more important than the other when you think about this school?” question was considered as it was proposing weighting as an obligation. So instead of this question, “Should there be an order of importance among these qualifications?” question was included in the schedule. Furthermore, the question “What do you think about the feasibility of this model? Would you like to be evaluated with such a model? Why?” was only letting the participant who wanted to be evaluated by this model explain their opinions. So that at the end of the question, the “Why not?” option is also added to the schedule.

Before the administration of the focus group interviews, the questions were piloted with one classroom teacher, one expert, and one assistant principal to see whether the questions were understandable and clear and whether the questions were working or not. After the piloting, some points in the explanation of the model like “Ensuring defensibility through face-to-face interviews” and “Evaluation model is open to updates” topics were expected to be explained in detail to the participants. The last version of the focus group discussion interview form is presented in Appendix D.

3.6 Data Collection Procedures

The researcher collected the data in the private school selected as the case during the academic years of 2018-2021. In the case studies, data collection procedures can follow a realist perspective aiming to collect data about events and behaviors or a relativist perspective aiming to collect data to capture the distinctive perspectives of the participants (Yin, 2018). As this research aims to reflect the different views of the participants for the development process of the model, the researcher preferred a relativist perspective. Also, in this research, since a model was developed with a

participatory approach, the participants' opinions were taken at every stage of the model development process.

3.6.1 One-to-one Interview

Data were collected from teachers, principals, and administrators through one-to-one interviews in the school setting. In case studies, interviews should be conducted in a friendly and non-threatening manner (Yin, 2018). Interviewing also includes observing the participants, and interviewers should consider how participants will respond if they seem uncomfortable (Maxwell, 2013). To propose a friendly and non-threatening environment and build a relationship, interviews could begin with a small talk made with participants about any topic from daily life or ice-breaking questions (Bogdan & Biklen, 2002; Creswell, 2015).

Before conducting interviews, the researcher obtained the necessary permission from the school's principal and the institution's general manager to conduct this research with teachers, experts, and principals. At the beginning of each interview session, the researcher informed participants that the principal and the school's general manager were aware of this research. This explanation made them feel comfortable instead of having an impression that they were conducting a secret or illegal process. As a result, they answered the questions sincerely. The researcher also made a short talk with the participants about their daily life experiences with questions like how they spent the day or about an event that has affected the school climate recently. Before each interview session, the researcher generated an interview environment where participants could feel comfortable. Besides, the researcher conducted the interviews in the form of a conversation and explained that these interviews were held for research purposes only to make participants feel comfortable and respond sincerely. In some cases, the researcher noticed that some participants did not want to explain the negative experiences. In this case, the researcher explained that they do not have to explain if they do not feel comfortable. As they trusted the researcher and the research itself, the participants started to talk about the processes they did not want to

talk about, making it possible to obtain in-depth and multifaceted data through the interviews.

The importance of their valuable contributions in the development of this teacher evaluation model was explained to the participants, and it was stated that thanks to these contributions, an effective model could be designed by deeply understanding the current situation in the school. The researcher also explained that the interviews were being performed just for research purposes. The researcher also stated that participants' names would not be used in any report or document.

The researcher herself conducted all the interview sessions to avoid data collector bias, and all of the interviews were conducted in quiet environments where participants felt comfortable. Interviews were audio-recorded after participants' permission. The researcher explained to participants that if they felt uncomfortable at any stage of the interview, they would terminate the interview session immediately, and the rest of the data would be deleted.

It is essential to ask questions naturally, maintain eye contact, and give feedback to the participants without sharing the researcher's ideas (Creswell, 2015). Questions asked by following a conversational approach and feedback to the interviewee were provided by nodding or using sentences like "I understand, thank you for sharing this experience with me," etc. All the participants were given enough time to answer the questions, or the researcher waited quietly while the interviewee was thinking. The researcher listened to the participants very carefully and took notes to avoid any problems while conducting the process or answering each question. At the end of each interview session, the researcher thanked the participants. Immediately after each interview, a short follow-up interview session was conducted in the follow-up sessions the notes taken by the researcher and the answers given to each question were directed the participant to clarify some points and give her the chance to make additions if she wanted to add further explanations. One-to-one interview sessions lasted forty minutes for teachers, thirty minutes for principals, and forty-five minutes for the experts.

3.6.2 Focus Group Interview

Focus group interviews work well when the participants feel comfortable, respected, listened to without being judged, and feel free to self-disclosure. In interview sessions, the researcher should create a permissive environment and control body language to encourage the participants to share their opinions and insights without feeling pressure (Krueger & Casey, 2015; Leedy & Ormrod, 2016). The permissions of each participant were obtained, and the purpose of the research conducted was explained before each focus group interview session. The researcher explained that the interview data would be used only to contribute to this research, and the confidentiality of the participants would be protected. Before the interview session, a daily topic was discussed with the participants in order to relieve possible tension.

It is crucial to arrange the focus group members because, in focus groups, the interaction between participants will yield the best information when the participants feel comfortable and cooperate (Creswell & Poth, 2018). Since the interaction of the participants with different experiences would be richer in focus groups, focus groups included participants who have contributed to the model's development process and who have not been interviewed before. While selecting participants in each group, the researcher considered the requests coming from the participants as well. For example, teachers stated that they would feel more comfortable with the assistant principal who works in their hall of classrooms and knows them. Therefore, the researcher assigned them to the same focus group. Four focus groups were formed. Each focus group included five teachers, one principal, and one expert.

During the focus group interview, the researcher, as a moderator, asked the questions, did not make any judgments, listened to the participants effectively, and made sure that everyone had an equal chance to talk. First, the processes conducted before, and the model developed so far were explained in detail. Then, participants were allowed to examine the model elements with the help of the model introduction document. Then, their opinions were gathered using the focus group interview form. The

participants were given a chance to explain their agreements or disagreements with each other.

The focus group discussions were held during the teachers' meeting hours, which is outside of the teachers' lesson hours, on the grounds that it may take longer than one class hour. Focus group interviews lasted between sixty-five and eighty-five minutes. They were audio-recorded after taking permission from all the participants. Since the interview room was not exceptionally large, the voice recorder was able to record everyone's voice. At the end of each interview session, the researcher thanked the participants. Immediately after the focus group discussion, a short follow-up interview session was conducted. In the follow-up sessions the notes taken by the researcher and the answers given to each question were directed to all participants to clarify some points and give them the chance to make additions if they wanted to add further explanations. Finally, participants were asked if they had any questions before ending the interview.

3.6.3 Online Feedback Forms

Participants' opinions regarding the guide and teacher evaluation tools were gathered through an online feedback form. The form, along with a detailed description of the research purpose and process, was sent to all classroom teachers, math and science teachers, principals, and experts working at the school via email. The form included the model and documents developed. The participants were asked to comment on them and provide any suggestions if necessary.

3.6.4 Pilot Testing of Teacher Evaluation Tools

In order to provide validity evidence for the teacher evaluation tools, the participants were asked to test the tools developed and to express their opinions on the clarity and evaluability of the expressions in these tools. Due to the Covid-19 epidemic, the practices for the course observation process were stopped at the school where the research was carried out. Therefore, only the self-evaluation tools were tested by

classroom teachers. Participants reported their opinions regarding the evaluation tools through email.

3.7 Role of the Researcher

One of the crucial features of qualitative research is positionality which refers to the researcher's role. In other words, social location in the context and settings like being a practitioner in the environment, being an expert, or being an outsider of the context (Ravitch & Mittenfelner, 2016). Researchers bring their background, such as work experiences, cultural experiences, and history, and their interpretations are not dependent on that background so that the readers need to know about the researcher. Reflexivity means reflecting on own experiences, background, and how the researcher viewed the process (Creswell, 2015). In qualitative studies, all the researcher's effort is used to understand the context of the phenomenon under the study, participants' beliefs, and behaviors in the study by social interaction (Patton, 2015). So that it is not appropriate to see social interaction as biased, and it is not possible to separate the interview from social interaction (Mason, 2002).

The researcher has been working as a curriculum development expert for five years in the school where the study was conducted. For five years, she worked with many of the classroom teachers working in the school collaboratively. The researcher developed lesson plans and teaching materials, designed materials to eliminate learning deficiencies and developed technology-supported teaching materials with the classroom teachers working at the school. Researcher also worked together on various projects. During these studies carried out with the teachers, the researcher had the opportunity to closely observe the qualifications of the teacher in important areas such as planning and preparation, instruction, monitoring students learning and reflective thinking. Moreover, researcher participated in the in-service trainings held in the school together with the teachers. In these trainings, researcher had the opportunity to closely examine teachers' perspectives on professional development.

During the years she worked in this institution, the researcher also worked in cooperation with the principals. The researcher, together with the principals took an

active role in the planning and conducting of the projects carried out for both students and teachers. She had the opportunity to understand the perspectives of principals on characteristics of effective teachers, teachers' professional development and teacher evaluation. In addition, as an expert, she worked with the experts working in the institution regularly and continuously. Thus, researcher had the opportunity to dominate experts' perspectives on professional development and teacher evaluation. The researcher's work in harmony with teachers, experts and principals and establishing positive relationships ensured that every participant took part in this research showed willing and sincere participation. There is a great need for an environment of trust between the researcher and the participants, especially in the study of a concept such as "teacher assessment", which is not welcomed by teachers. The fact that the researcher works with the teachers, principals and experts working in the school in harmony, by establishing positive relationships and with devotion made it possible to study the subject of this research easily and effectively.

3.8 Data Analysis

In qualitative research, emphasis is on description rather than using numbers to determine the relations (Maxwell, 2013). The data analysis process includes arranging the interview transcripts or fieldnotes systematically, organizing data, breaking them into manageable units, coding them, synthesizing them, and searching for patterns to enable the researcher to come up with the findings (Bogdan & Biklen, 2007). Qualitative researchers can follow inductive or deductive way by using their reasoning skills. Creating codes, themes, and categories inductively means organizing the data from bottom to top (Creswell & Poth, 2018). In the present study, the analysis was done through content analysis; codes, themes, and categories were created inductively. At the beginning of the data analysis process, the collected data were transcribed by the researcher and made ready for the analysis. The data analysis process of the one-to-one interviews and focus group interviews conducted within the scope of this research was carried out in a similar way.

3.8.1 Preparation for the Data Analysis Process

Making sense of a massive amount of data is the challenge of qualitative analysis, and this process involves identifying significant patterns and constructing a framework (Patton, 2015). The first step of content analysis is reading the interview transcripts written by listening to the interview recordings. The researcher transcribed the audiotapes by listening to the recordings and typing word by word using the Microsoft word processing program. The transcribed text was read by the researcher line by line. This process helped the researcher be familiar with the transcribed data and remember most of the interview content while coding and analyzing the relation between codes to produce themes and categories. During the data collection phase, the researcher noted the general impressions and observations obtained during interviews.

3.8.2 Coding and Constructing Categories

In qualitative research, code is the form of words or phrases that express the salient, essence-capturing and/or evocative qualities of the data, which is usually based on language or visual elements (Saldaña, 2016). In other words, coding can be defined as the process of using words or phrases to reveal the situation in the data pieces obtained after data collection and to define the meaning specified in these data pieces (Creswell, 2013; Rossman & Rallis, 2012). The obtained codes are used to associate the data with conceptual frameworks and more comprehensive concepts (McAlister et al., 2017). There are two processes that are frequently used in coding, which are open coding or using predefined codes (Creswell, 2014; Patton, 2015; Saldaña, 2016). In open coding, in other words, inductive coding, while analyzing the data, the researcher develops codes based on the conceptual framework or the investigated phenomenon. On the other hand, while using predefined codes, the researcher can develop codes based on the conceptual framework before the analysis process and tries to find these codes in the data (Creswell, 2014; Patton, 2015). In this study, many different views and possibilities were needed on the grounds that a teacher evaluation model, which cannot be included in the current situation, would be developed with a model-specific to the school.

For this purpose, for both one-to-one and focus group interviews, the open coding process was carried out. Coding is a stage where the researcher begins to think deeply about the data, develop familiarity with the data, catch certain nuances, and even shape some categories (Saldaña, 2016). According to Miles et al. (2014), coding includes two different cycles. The first cycle of coding is about summarizing and condensation the data into readily analyzable units. On the other hand, the second cycle of coding is about pattern coding used to group those summaries into categories, themes, or explanatory and more meaningful constructs. While determining the themes and categories, reviewing the literature might be influential (Bogdan & Biklen, 2007).

The researcher coded transcribed audio of the individual interviews recordings manually. The transcribed text was analyzed for the first coding cycle, and words or sentences that make sense were determined. The researcher noted the codes on the right-side margin of the transcribed data. In the second cycle of coding, the relationship between the codes was examined, and codes representing a similar dimension were categorized under the same categories. While developing the categories and themes, the researcher considered the aim, research questions, and theoretical framework, including teacher evaluation models, teacher evaluation frameworks, and related literature research. Some themes and categories are continuously changed when the researcher found better and more descriptive phrases and recoding the data several times, contributing to the study's trustworthiness. An example from the researcher's notes is given in Table 3.4 to set an example for the coding and developing categories process of face-to-face interviews.

Part of Participant T1's Responses, Codes, and Categories

In qualitative research, it is suggested to develop a codebook to prevent the codes from being unclear or meaningless and to overcome the meaning shifts that may occur while developing the codes (Creswell, 2014; Miles & Huberman, 1994). The detailed descriptions of the code, constraints and concrete examples of each code are given in a codebook to aid the analysis of large sets of qualitative data (Roberts et al., 2019). Following the recommendation, the researcher created a codebook (including code, its description, and sample quotations) for one-to-one interviews (see Appendix E) and for focus group interviews (see Appendix F).

3.8.3 Reporting of the Findings

In the reporting process of qualitative data, it is essential to include rich, detailed descriptions and direct quotations to allow the reader to understand the participants' thoughts represented in the report and to provide evidence to illuminate the case (Patton, 2015). Therefore, the researcher highlighted the significant statements by quoting almost all the findings reported under essential topics. In the reporting process of the findings, the researcher took those codes, themes, and categories as the basis.

3.9 Trustworthiness of the Study

Reasonable standards that help researchers judge the quality of the conclusions from the research findings can be referred to as the quality of the research (Miles & Huberman, 1994). In qualitative studies, reliability, and validity, also referred to as trustworthiness, are vital components, and achieving rigor leads to methodological validity (Lincoln & Guba, 1985; Ravitch & Mittenfelner, 2016). Trustworthiness of the research is discussed under four headings; "Credibility (Internal Validity)," "Transferability (External Validity/ Generalizability)," "Dependability (Internal Reliability)," and "Confirmability (External Reliability)."

3.9.1 Credibility

Credibility is about how congruent the findings are with the reality; in other words, how the findings are true and accurate. Triangulation which includes collecting information from various individuals or settings by using multiple methods, reduces the risk of chance associations and biases (Maxwell, 2013). The data collected in this research reflect participants' views from different perspectives working in various fields (teachers, principals, and experts) in this school. From this point of view, the researcher provided the triangulation of the resources. On the other hand, to avoid the inaccuracy of the data, the researcher recorded all interviews and used verbatim audio transcription of these recordings.

For prolonged engagement, it is essential to spend enough time interviewing, learning the culture, building trust and sound relationships with respondents, and being frank and comprehensive about what the participants tell (Patton, 2015). The researcher has been working at the school for many years, and during this research, she spent enough time with all the participants. The fact that the researcher knew the school's culture and previously worked with most of the participants allowed the participants to trust the researcher and reflect the statements they made as reflecting the reality. Furthermore, to collect accurate and relevant data, the researcher allocated sufficient time to each participant while interviewing and searching for alternative or opposite explanations.

Peer debriefing and member checking are used to ensure that the researcher's conclusions are grounded in the data (Creswell & Poth, 2018). Among other procedures, peer debriefing and member check can be defined as the most appropriate methods for credibility (Patton, 2015). Peer debriefing meaning that consulting an expert or mentor on data collection, data analysis, and reporting process of the research to ensure whether this process is conducted by following an objective approach and challenge the researcher on the assumptions, biases, and interpretations at many stages of a research study (Ravitch & Mittenfelner, 2016). In this study, the researcher consulted the Dissertation Committee Members, the researcher's thesis advisor, and two researchers experienced in qualitative research studies during each research phase to gather feedback.

Member checking and member reflections refer to the occasions that allow for sharing and dialoguing with participants about the findings of the study, interpretations of the researcher by providing critique or feedback from them, and finally, including participants' points of view in data analysis phases (Creswell, 2015; Tracy, 2020). Member checking is the most important way of ruling out the possibility of misinterpreting the responses' meaning, and the participants' perspectives have usually been used for respondent validation (Maxwell, 2013). On the other hand, member reflections suggest that participant feedback is valuable as a measure of credibility and for additional insight gathered through collaboration and elaboration

(Tracy, 2020). After each interview, the researcher summarized the answers to the interviewed participants to get validation, correct the misunderstandings, or offer additional information. Furthermore, the teacher evaluation model based on the data gathered from individual interviews was checked by the participants through focus group interviews led by the researcher.

3.9.2 Transferability

Transferability is different from generalizations because for generalization researcher engages in random sampling and objective scientific practices to generate context-free generalizable knowledge, but in qualitative studies, the researcher defines the context in detail and communicates the impact of the finding to the reader to make them imagine and personally transfer these findings to familiar settings (Tracy, 2020). Transferability dealt with the issue of generalization in case-to-case transfer by providing readers with sufficient information on the case studied so that the reader could establish the degree of similarity between the case studied and the case in which findings might be transferred (Patton, 2015). In this study, for the findings obtained from this research to apply to another case with a similar context, a thick description of the case was given in detail under the heading “Context of the study.” While reporting the study’s findings, the researcher also used quotations for the readers to understand the context and the participants’ feelings and perceptions. On the other hand, purposive sampling was employed while selecting the cases and participants to maximize the range of specific information obtained from that context. Other researchers in the field might transfer the findings of this research to other cases owing to some common features.

3.9.3 Dependability

In quantitative studies in which positivist techniques are employed, reliability is addressed if the work is repeated in the same context, with the same methods, and with the same participants, researchers will obtain similar results. In a qualitative study, the dependability term is used instead of internal reliability. A future researcher could

repeat the work if the process within the study is reported in detail, which also allows the reader to assess whether the research practices have been followed adequately or not (Shenton, 2004). To provide dependability, an external audit, like members of the doctoral committee, can render judgment about the quality of data collection and data analysis process (Patton, 2015). In this research, both the data collection process and the data analysis process were examined in detail by the researcher's thesis advisor, members of the doctoral committee, and two researchers working in the field of qualitative research. The researcher shaped the research process and sometimes carried out the procedures according to these people's feedback and valuable contributions.

3.9.4 Confirmability

Confirmability, which is also considered objectivity, is about proofing that the data and interpretations of an inquiry are not figments of the researcher's imagination (Patton, 2015). Confirmability is about keeping researcher bias at a minimum level and making the conclusions depend on the subject and conditions of the inquiry rather than on the inquirer research (Miles & Huberman, 1994). It is about ensuring that data generation and analysis have been appropriate to research questions and done carefully, honestly, and accurately (Mason, 2002). One way to ensure confirmability is to examine the reports in which each stage of the study is explained in detail by an audit trail, that is, a reliable outside observer (Lincoln & Guba, 1985). Continuous feedback was received from the thesis advisor on important issues such as whether the general methods and procedures of this research were explained clearly and in detail, how the data were collected, processed, and transformed, and whether the findings were clearly related to the research.

The researcher should be as open and self-aware as possible about personal assumptions, values and prejudices, and emotional states and how these may have come into play during the study (Miles et al., 2014). For this purpose, notes in the form of reflective explanations can be added during the research process (Ortlipp, 2015). To provide confirmability, the researcher kept notes after each day in which interviews

were conducted and recorded initial impressions of each data collection session, potential biases, and predispositions that may affect the research process and conclusions. The researcher benefited from these notes, especially in the data analysis process and reporting results section. An example of one of the notes kept by the researcher is given below:

Some of the teachers working at the school, especially the department heads, are very open to evaluation. This openness may be due to their assuming the identity of a principal because the school principal asks them for opinions about the other teachers. They are willing to fulfill this responsibility more accurately with evaluation. It can be said that young teachers at that school are also quite open to teacher evaluation. Young (less experienced) teachers seem overwhelmed by different practices applied to experienced teachers because they say that these applications are made to favor experienced teachers. I understand this situation because they constantly complain during the interview and want them to be treated equally with experienced teachers in this new evaluation process. They look very positive at the idea that experienced teachers will be evaluated under the same conditions as themselves. I am surprised that the teachers I never expected were so open to evaluation. If the teacher believes that this work is done correctly, it seems that she will accept the evaluation model.

3.10 Limitations of the Study

Although unique cases can provide valuable evidence, the results of this study are limited to those cases and settings. Even if the case in which this research was conducted was described in detail, it would not be correct to use the results of this research as it is in another institution or make predictions for another school or institution. However, it is thought that the way the research is conducted, and the way followed while reaching the results can be a guide for institutions or individuals who want to design an evaluation model specific to a school and meeting the needs of the people working in that school.

Due to the Covid-19 pandemic process that took place in the years when the study was continuing, the classroom observation processes carried out by someone other than the teacher were stopped due to the online lessons and the health precautions taken in the classroom after the pandemic process. Therefore, the piloting process of classroom observation forms developed within the scope of this model could not be fully realized.

All teachers, experts, and principals participating in the research stated that they are eager to implement this evaluation model. They participated and contributed to the development of the model. However, the school principals believed that implementing this model would not be realized immediately. It will take time to make the necessary preparations, and the school can implement them in the coming years. So that the effectiveness of the model could not be evaluated because the model could not be applied.



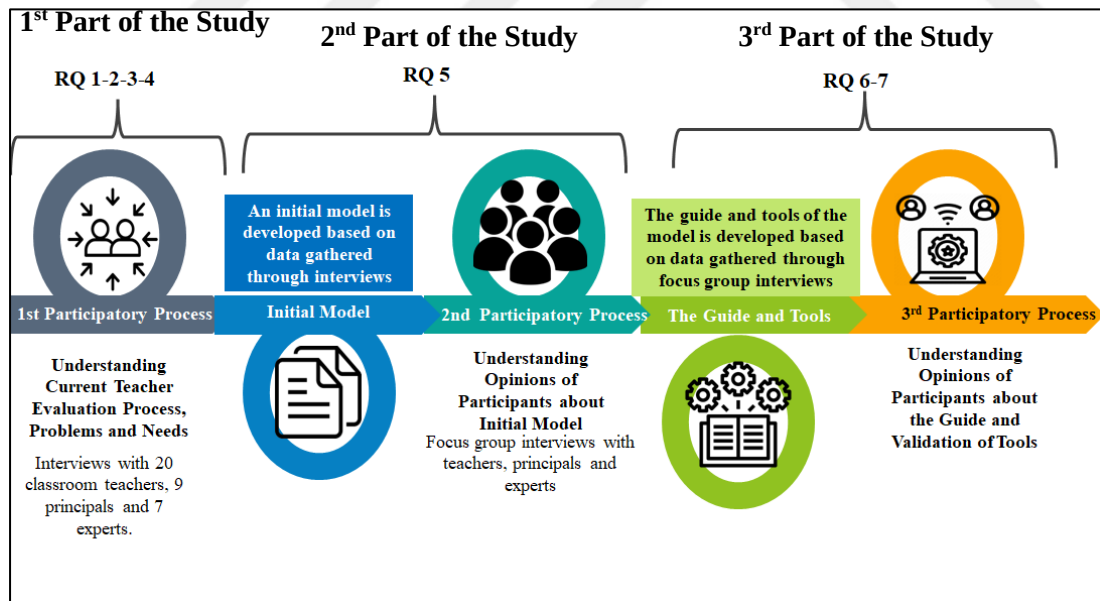
CHAPTER 4

FINDINGS

This chapter presents the results of the study in line with the research questions. Since this study was conducted in three different participatory processes, as explained in the design (Figure 3.1), the findings are presented in three parts. The first four research questions were addressed in the first part, the fifth research question in the second part, and the sixth and seventh research questions in the third part.

Figure 3.1

The overall design of the study



4.1 First Part of The Study

The first part of the research aimed to provide an in-depth analysis of perceptions of classroom teachers, principals, and experts on current teacher evaluation and

professional development practices at the private school and to examine the suggestions of the participants in depth. For the first part of the research, data were collected through one-on-one interviews, and data were analyzed with content analysis. The themes, sub-themes, and the codes that emerged as a result of the data analysis are given in tables.

4.1.1 Teacher Evaluation Practices Carried Out (Research Question 1)

To answer the first research question regarding the current state of the school's teacher evaluation process, the researcher gathered data through interviews conducted with teachers, principals, and experts working in that school. The relevant codes provided by the content analysis were organized under the themes: "evaluation process" and "use of evaluation data."

4.1.1.1 Evaluation Process

The codes of the statements that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the current teacher evaluation process carried out at the school are given in Table 4.1.

Table 4.1

Codes for The Evaluation Process

<i>Theme 1. Evaluation Process</i>
No evaluation schedule or pattern
No standardized forms
No written purpose for evaluation
No defined standards or criteria
Data source: Average achievement scores
Data source: Lesson plans
Data source: Informal principal observation
Data source: Opinions of parents
Data source: Opinions of colleagues

Findings related to this theme showed that the majority of teachers ($f=18$) indicated no specific schedule or timeline for evaluation. Only two teachers (T6, T12) mentioned

that after each school-wide exam, which is only based on academic achievement scores, teachers received feedback from principals. This feedback process was considered a pattern for the evaluation process. Similarly, all the principals and experts also stated no specific schedule or timeline for evaluation.

All of the principals, most experts ($f=6$), and most teachers ($f=17$) mentioned that standardized forms are not being used in this school. Instead, teachers stated that informal principal observations are being conducted for the first time in the 2018-2019 academic year. Only three (T4, T6, T7) of them stated that they did not know whether any form was being used. In addition, the school principal (P4) also indicated that she did not use any form during the observations.

Most of the teachers ($f=17$) stated that the purpose of the evaluation was not determined clearly. Most of the experts ($f=5$) also expressed their opinion that there was no purpose in a written form. On the other hand, three experts (E1, E2, E6) stated that even though it was not written, it was evident that evaluation was made to increase student achievement in this school. On the other hand, some of the principals ($f=4$) mentioned that the rationale for the evaluation is both for professional development and better student achievement.

One of the teachers explained this situation as:

I do not know clearly, what the evaluation was made for, and suppose the evaluation was conducted based on a purpose. Unfortunately, no one informed us about this purpose, written or orally. We are just trying to have ideas about the aim of the evaluations by making predictions (T1 four years of experience).

Değerlendirmenin ne için yapıldığını net olarak bilmiyorum ya da değerlendirmeler bir amaca dayalı olarak yapılıyorsa da ne yazık ki kimse bu amaç hakkında bizi şimdiye kadar yazılı ya da sözlü bilgilendirmedi. Aslında, sadece tahminler yaparak değerlendirmelerin amacı hakkında fikir sahibi olmaya çalışıyoruz. (T1 dört yıllık tecrübe)

Findings revealed that all the teachers thought there were no criteria explained before or after the evaluation process, so they were unaware of how they were being evaluated. Similarly, most principals ($f=8$) and experts ($f=6$) stated that there were no defined standards or criteria used in the teacher evaluation process. An assistant

principal (P2) and an expert (E7) stated that they were not sure whether criteria were used or not during the observation process. One of the teachers explained as:

This year our school principal did the observation, but I am not aware of how I am evaluated. During the observation process, she took notes, but it is not clear that she wrote these notes under what headings, and I do not know what criteria were used. I don't even think any criteria were used. (T6 six years of experience)

Bu yıl okul müdürümüz gözlem yaptı, ancak nasıl değerlendirildiğimi bilmiyorum. Gözlem süresince notlar aldı, bu notların hangi başlıklar altında alındığı açık değil ve hangi ölçütleri kullandı bilmiyorum. Hatta bence ölçüt bile kullanılmadı. (T6 altı yıllık tecrübe)

When participants were asked how the information obtaining process for evaluation is being carried out in the school, all participants stated that scores gathered from school-wide exams, which are only based on academic achievement used as the data source for teacher evaluation. When they were asked how these scores were used to evaluate teachers, they stated that after each school-wide exam, a report that includes an average score of students in a class was presented to all the teachers. Based on these reports, the success of the teachers' classrooms was regarded as equivalent to the teacher's success. Similarly, all experts and some principals (P2, P3, P5, P7, P8) explained the use of achievement scores in the teacher evaluation process similar to the teachers. In addition to this view, some of the principals (P1, P4, P6, P9) in which the school principal takes part also added that the average score is essential, but rather than applying as a sanction, average scores were used to understand the causes of failure. An excerpt from the interview with the school principal is as follows:

Of course, exam results are essential, and they explain to me what the teacher did in that class, but without putting pressure on the teacher. For example, I took the exam results this year and said, "please explain the situation of these students in an objective way" to the teachers. I'm going to thank you even if you told me that you did not do anything for this student. I need to know the truth to understand this failure". (P4 seventeen years of experience).

Sınav sonuçları önemli, öğretmenin o sınıfta neler yaptığını bana açıklıyor ama öğretmen üstünde baskı kurmadan. Örneğin, bu yıl sınavların sonuçlarını önüme aldım ve öğretmenlere "lütfen bu öğrencilerin durumunu objektif bir şekilde bana açıklayın" dedim. Bana bu öğrenci için hiçbir şey yapmadığınızı söylemiş olsanız bile size teşekkür edeceğim. Bu başarısızlığı anlamak için gerçeği bilmem gerekiyor" diyorum. (P4 on yedi yıllık tecrübe)

All the teachers stated that they report their lesson plans to the assistant principal each week, and these lesson plans were also used for obtaining data for teacher evaluation. Among the experts, only four experts (E1, E4, E5, E7) stated that teachers' lesson plans were evaluated. Other experts, on the other hand, stated that they were not aware of the process of sending the lesson plans to the assistant principal and the evaluation process of these plans. Most of the principals ($f=8$) stated that they were using lesson plans to provide evidence for measuring teachers' effectiveness. Three assistant principals (P2, P3, P6) said they used lesson plans to understand how well teachers prepared for the course and how well they planned it. Five of them (P1, P3, P4, P5, P8, P9) stated that lesson plans were used in the evaluation process because they contain clear information about how the lesson will be carried out and provide evidence about how effectively the teacher will conduct the lesson. One assistant principal explained the use of lesson plans as a resource for evaluation as follows:

We evaluate lesson plans because we use them in every aspect of life, and I think making a good plan is half the success. Teachers are generally lazy. Especially those who are experienced, "do I have to make a plan at this age?" says. I think a good plan means that it is well planned and organized in the lesson. In addition, the more successfully a teacher plans the lesson, the more effective and successful it will be in the lesson. (P3 sixteen years of experience).

Ders planlarını değerlendiriyoruz çünkü planları hayatın her alanında kullanıyoruz ve bence iyi bir plan yapmak başarının yarısı demektir. Öğretmenler, genelde üşeniyor. Özellikle deneyimli olanlar "bu yaşta plan yapmak zorunda mıyım?" diyor. İyi plan dersinde iyi planlanması ve organize edilmesi demek bence. Ayrıca bir öğretmen dersi ne kadar başarılı planlarsa, derste o kadar etkili ve başarılı olur. (P3 on altı yıllık tecrübe)

All the principals, almost all the teachers ($f=18$), and experts ($f=7$) also stated that the principal conducted informal observations in unannounced way for the first time in 2018. Teachers noted that the school principal sat at the back of the class in this observation process and kept notes. Two of the teachers (T13, T16) indicated that they had not yet been observed, and they also stated that the school principal would not have had enough time to observe their classes.

Some teachers (T2, T6, T18) and experts (E1, E4, E5) stated that even though there was no legal implication, parent opinions were also considered an essential indicator

for the teacher evaluation process. Furthermore, they explained that the parents' views were taken into consideration because this institution is private, and sometimes the teachers were warned about the parent opinions. Some teachers (T4, T9, T19) stated that the views of the head of the department, who are experienced teachers at the school, were taken into consideration, and the opinions of these people were sought, especially about the teachers who have just started working at the school.

4.1.1.2 Use of Evaluation Data

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts on how the data were obtained through the current teacher evaluation process in the school are given in Table 4.2.

Table 4.2

Codes for The Use of Evaluation Data

<i>Theme 2. Use of Evaluation Data</i>
Improving student achievement
Predicting teachers' effectiveness
Improving the quality of school
Providing summative feedback
Providing individual formative feedback

Participants were asked about how evaluation results were being used. Most of the teachers ($f=18$), principals ($f=7$), and all the experts stated that although the school did not have an announced evaluation rationale, the list of results that emerged after the general exams was used to improve the academic achievement of the students. Half of the teachers, most of the principals ($f=6$), and experts ($f=4$) mentioned that these results were also used to predict the effectiveness of the teacher like good, bad, successful, failed, etc. They also emphasized that the teacher with an excellent average grade was considered successful, and the teacher with a low average score was considered unsuccessful. Some of the principals (P2, P7, P8, P9) also stated that teacher evaluation was generally conducted to improve the quality of the school. When participants were asked how they have taken feedback, all the teachers, principals, and

experts stated that principals provided summative feedback each term by conducting meetings attended by all teachers who taught the same grade level. In addition, three teachers (T2, T10, T12) and four principals (P1, P4, P5, P7) mentioned that principals provided individual feedback mainly when problematic situations occur.

4.1.2 Perceived Strength and Weaknesses (Research Question 2)

To answer the second research question regarding the strength and weaknesses of the teacher evaluation process, the researcher gathered data through interviews conducted with teachers, principals, and experts working in that school. When the situation, in this case, was examined, the participants mentioned weaknesses rather than the strengths of the current evaluation process. Therefore, under this research question, the problems experienced in teacher evaluation have been mentioned. The relevant codes provided by the content analysis were organized under propriety, utility, and accuracy.

4.1.2.1 Propriety Problems

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the propriety problems of the current teacher evaluation process in the school are given in Table 4.3.

Table 4.3

Codes for Propriety Problems

<i>Theme 1. Propriety</i>
Not providing policies and procedures
No access to evaluation results
No balanced evaluation

While ensuring the propriety, evaluations should be conducted legally and ethically, and the welfare of the participants involved in the evaluation should be considered (Howard & Gullickson, 2009). Policy statements and/or guides that describe the purpose of the evaluation system, how the system will be used, and the data collection

and reporting process, ensure that assessments are consistent, fair, and equitable. Most of the teachers ($f=18$), principals ($f=8$), and all the experts stated that there was no guide, policy, or procedure in which the timing, rationale, or schedule are explained. Teachers explained this situation raised doubts about being equitable, providing fairness, and consistency of the evaluations made. Principals and experts also stated that this situation caused teachers to worry about the propriety of the evaluation process, and therefore they always felt uneasy. One of the teachers (T3 *fourteen years of experience*) explained, *"The purpose and the process are unclear. This brings the question of why I'm being evaluated. I also do not know whether this evaluation is special to me or if it is done for everyone in an equal manner. (Amaç ve süreç net değil. Bu ben neden değerlendiriliyorum sorusunu sormama neden oluyor. Bu değerlendirmenin bana özel olup olmadığını veya herkese eşit şekilde yapılıp yapılmadığını da bilmiyorum."* Another teacher (T6 *six years of experience*) expressed the problem caused by the uncertainty of how often the evaluation was made as follows. *"The lack of a certain frequency of observation means that everyone should be as comfortable as possible. All the teachers in this school are very comfortable and does not need to improve themselves (Belirli bir gözlem sıklığının olmaması, herkesin olabildiğince rahat olmasın sağlıyor. Bu okuldaki tüm öğretmenler çok rahat ve kendilerini geliştirmeleri gerekmiyor)"*. One of the experts (E3 *ten years of experience*) stated as *"It is unclear whether enough time is allowed for the teachers to change and correct themselves because the schedule is unclear. In this case, it is necessary to question whether the work done is correct or ethical. (Öğretmenlerin kendilerini değiştirmeleri ve düzeltmeleri için yeterli zamanın verilip verilmediği belli değil çünkü zaman çizelgesi belirsiz. Bu durumda yapılan iş doğru ya da etik mi sorgulamak lazım.)"* A principal expressed as follows.

In the corridors, I sometimes try to make observations. Passing through the hallway, I peek through the door to the teacher's class. Try to get an idea of what I can see at that moment. But I don't know how often I should observe as an administrator. I don't know which qualification I should observe. It is not possible to make an

Koridorlarda bazen gözlem yapmaya çalışıyorum. Koridordan geçerek öğretmenin sınıfına kapıdan bakıyorum. O anda ne görebildiysem bir fikir edinmeye çalışıyorum. Ama aslında, yönetici olarak ne sıklıkla gözlemlemem gerektiğini bilmiyorum. Hangi niteliklere bakmam gerektiğini bilmiyorum. Sadece sınıfı kapısından

accurate or fair assessment with only two minutes of observation from the classroom door. It's not ethical either. (P2 twenty-two years of experience). *yapılan iki dakikalık gözlemle doğru ya da adil değerlendirme yapmak mümkün değil ki. Bi kere etik de değil. (P2 yirmi iki yıllık tecrübe)*

A great majority of the teachers ($f=15$) stated that they were not informed before or after the observations; besides, they never get individual feedback from principals about the evaluation results. Therefore, they indicated that they were not aware of the weaknesses or strengths that have been observed or evaluated. Teachers felt uncomfortable because they thought that the information gathered through classroom observations or from any other evaluation that they were not aware of was being used for some other reasons. A teacher (T11 twenty years of experience) explained this problem with the following statement: *"I don't know the result of the evaluation at all. Feedback is given in general meetings. However, the feedback should not be general. One must know what is right and what is wrong and how to improve herself. (Değerlendirme sonuçlarını hiç bilmiyorum. Genel toplantılarda geri bildirimler veriliyor. Geri bildirim genel olmamalıdır. Kişi neyin doğru neyin yanlış olduğunu ve kendini nasıl geliştireceğini bilmelidir.)"*

Most of the teachers ($f=12$) stated that their feedback was not balanced, and generally, they received individual feedback about their weaknesses. The teachers indicated that they were received one-to-one feedback, especially when their class average achievement scores decrease in school-wide exams, when there were problems experienced throughout the school and when parents complained about something that concerns them. Few teachers (T5, T11, T16) stated that they are generally praised and appreciated for working in this school for many years. In general, experienced teachers noted that the principals appreciate their work, which is expressed face to face. One of the experienced teachers explained this situation as:

I am an experienced teacher, and principals always appreciate me, but most teachers worry when a principal call her and asks to meet with her because principals only talk about negatives when they want to speak individually. I *Deneyimli bir öğretmenim ve müdürüm beni her zaman takdir eder, ancak müdür bir başka öğretmeni aradığında ve onunla görüşmeyi istediğinde öğretmenlerin çoğu endişeleniyor çünkü yöneticiler bireysel olarak konuşmak istediklerinde sadece olumsuz gördüğü*

wish teachers were also told so often about the things they did well (T11 twenty years of experience). şeyleri söylüyor. Keşke öğretmenlere iyi yaptıkları şeyler hakkında da sık sık bilgi verilse (T11 yirmi yıllık tecrübe).

While some of the principals ($f=5$) and most of the experts ($f=5$) shared the same opinion with the teachers, a group of principals (P1, P4), including the school principal, stated that they frequently gave one-to-one feedback to the teachers and that these feedbacks were both positive and negative.

I give direct feedback. These feedbacks are beneficial for both encouraging the teacher and identifying the deficiencies. For example, there was a well-done work in a classroom where I observed. After the observation, I immediately appreciated the teacher and asked her to share it with other teachers. I even told the head of the group that she should encourage this teacher. (P4 seventeen years of experience). Birebir dönütlerde veriyorum. Bu dönütler hem öğretmeni yüreklendirmek hem de eksiklikleri belirlemek için çok faydalı. Örneğin gözlem yaptığım bir sınıfta güzel yapılan bir çalışma vardı gözlem sonrası öğretmeni hemen takdir ettim ve bunu diğer öğretmenlerle de paylaşmasını istedim. Hatta zümre başkanına bu öğretmeni yüreklendirmesi gerektiğini söyledim. (P4 onyedili yıllık tecrübe)

Furthermore, a few principals (P6, P7) and some experts (E5, E7) stated that they did not know how the feedback processes were carried out.

4.1.2.2 Utility Problems

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the utility problems of the current teacher evaluation process in the school are given in Table 4.4.

Table 4.4

Codes for Utility Problems

Theme 2. Utility
Lack of explicit criteria
Lack of functional reporting

Utility term is used as a degree of being useful, and for ensuring utility, evaluations should be informative, timely, and influential (Howard and Gullickson, 2009). For the usefulness of the results, evaluation systems must outline expectations for performance by using well-defined and explicit criteria. In evaluations, well-defined criteria should be used to interpret or judge the performance based on a clear and defensible rationale. Otherwise, individual interpretations take their place, and these interpretations jeopardize the usefulness of the results. Almost all the teachers ($f=17$) mentioned the lack of explicit criteria as a problem and stated that defensibility in evaluation was essential and can be provided by setting explicit criteria. One of the teachers described the interpretation and judgment process problem as:

Principals have opinions about me, but I do not know how they have achieved this view. I don't know which performance indicator or criteria was used. Similarly, principals have a negative opinion about a teacher who has just started school. She also does not know how this idea came into being. I can neither defend myself nor her. (T11 twenty years of experience)

Yöneticilerin benim hakkımda görüşleri var ama bu görüşe nasıl ulaştıklarını bilmiyorum. Hangi performans göstergesi ya da ölçüt kullanıldı bilmiyorum. Benzer şekilde, yöneticilerin okula yeni başlayan bir öğretmen hakkında da olumsuz bir görüşü vardır. Bu fikrin nasıl ortaya çıktığını bu yeni öğretmen de bilmiyor. Ne ben kendimi savunabilirim ne de o. (T11 onbir yıllık tecrübe)

Another teacher (T18 *eighteen years of experience*) also stated as, “I must know what is measured in this observation, what are the criteria. I think it is meaningless and not useful when the evaluation is not done with the criteria. (Bu gözlemden neyin ölçüldüğünü, ölçütlerin neler olduğunu bilmeliyim. Değerlendirmenin ölçütlerle yapılmadığında anlamsız ve yararsız olduğunu düşünüyorum.)” Similarly, most principals ($f=7$) mentioned no specific criteria used for interpretation or judgments in the evaluation process. Some of the principals (P1, P3, P6, P8) explained that they were using their point of view more than written criteria while making judgments. One of the principals explained this process as:

So, I often walk around my hall and do observations from the classroom door. Even this kind of observation allows you to make some judgments about the

Genellikle koridorumda dolaşıyorum ve sınıf kapısından gözlemler yapıyorum. İnan bana, ölçüt olmasa da bu tür bir gözlem bile öğretmen hakkında bazı kararlar vermemi

teacher even though there are no criteria. Because I was a teacher, and I know what effective teaching means. (P8 seven years of experience). *sağladı. Çünkü ben de bir öğretmenim ve etkili öğretimin ne anlama geldiğini biliyorum. (P8 yedi yıllık tecrübe).*

All the experts stated that they did not think that a criterion is used. Some of them (E1, E2, E4, E6) noted that the principals received support from experts in creating standards for the different forms to be applied in the school, but they had never been consulted in developing criteria for teacher evaluation. All the experts stated that objectivity was not possible if there are no explicit criteria. One of the experts (E5 two years of experience) explained this situation “If there are no criteria, an objective result cannot be obtained, and generalization of the teacher’s performance cannot be made. (Ölçüt yoksa, nesnel bir sonuç elde edilemez ve öğretmenin performansı hakkında genelleme yapılamaz.)” Another one expressed the problem caused by the lack of criteria as follows:

To be fair and equitable, we explain when students pass or fail an exam with indicators and reasons. We also give them the opportunity to express themselves. But here I am not given the right to speak [...] This is perhaps the most important problem to be solved for the school. One solution is to set certain criteria so that every teacher knows why the decisions were made and can have an opportunity to defend herself. The explanation is made according to certain criteria, it becomes more meaningful, and this evaluation process is the same for all teachers. In other words, if it is applied without discrimination among teachers, I accept this evaluation. (T2 five years of experience). *Eşit ve adil olmak için öğrenciler bir sınavda başarılı ya da başarısız olurlarsa bunu göstergelerle ve nedenleriyle açıklıyoruz. Onlara kendilerini ifade etme fırsatı da veriyoruz aslında. Fakat burada bana konuşma hakkı verilmiyor [...] Bu, okul için çözülmesi gereken belki de en önemli sorun. Çözümlerden biri belirli ölçütler belirlemek, böylece her öğretmen alınan kararların neden alındığını bilir ve kendini savunma fırsatı bulabilir. Açıklama belirli ölçütlere göre yapılır, daha anlamlı olur ve bu değerlendirme süreci tüm öğretmenler için aynı olur. Yani öğretmenler arasında ayırım gözetmeksizin uygulanırsa bu değerlendirmeyi kabul ederim. (T2 beş yıllık tecrübe).*

One expert expressed the problems caused by the lack of criteria as follows:

As long as there are no criteria, it means that we are using our *Ölçüt olmadığı sürece, kendi yargılarımızı kullanıyoruz demektir.*

judgments. How can we use our subjective judgments as to the evaluation result, and how can we define the things to be improved because of the evaluation? Then we should ask ourselves why we do the evaluation. (E3 *ten years of experience*).
Öznel yargılarımızı nasıl kullanabiliriz ve değerlendirme sonucunda iyileştirilecek şeyleri nasıl tanımlayabiliriz. Aslında, o zaman değerlendirmeyi neden yaptığımızı kendimize sormalıyız. (E3 *on yıllık tecrübe*)

For the evaluations to be effective, all users should understand the results and actions included in a functional report. In addition, reported results should be given regularly and explained to the teachers to pursue appropriate actions (Howard & Gullickson, 2009). For this study, another critical problem for utility is defined as a lack of functional reporting. Most of the teachers ($f=16$), principals ($f=7$), and all the experts reflected that there was no written report written in a timely manner. Participants defined this as a problem because, according to them, reporting was essential to provide further development of strengths and improvement of weaknesses. Furthermore, reporting that includes feedback for professional development allows the teachers to develop a plan, and the reporting process will help keep the development under record. One of the teachers (T1 *four years of experience*) defined this problem as “*If one had checked me and kept the report, it would keep me vigorous. In this way, I know my weaknesses and monitor my progress. (Biri beni kontrol etse ve rapor tutsa, aslında bu beni dinç tutar. Bu şekilde zayıf yönlerimi bilirim ve gelişmeleri takip ederim.)*” One teacher reflected on this problem as,

Now, if someone asks me about what kind of training I need to attend, I can only explain according to my self-evaluation. I could give a more accurate answer to this question if the feedback on my weaknesses were presented at regular intervals within a written report. (T8 *eighteen years of experience*)
Şimdi, birisi bana ne tür eğitimlere ihtiyaç duyduğumu sorsa, sadece kendi değerlendirmem doğrultusunda açıklayabilirim. Zayıf yönlerimle ilgili geri bildirimler yazılı bir raporla düzenli aralıklarla sunulsaydı bu soruya daha doğru yanıtlarım. (T8 *onsekiz yıllık tecrübe*)

Few teachers (T7, T13, T20) stated that they did not see the lack of a reporting process as a problem. They have indicated that the oral feedback they have received is enough. One of them (T13 *twelve years of experience*) explained her opinion: “*The lack of an assessment on paper is not very necessary. We already know each other very well*

because we have been working together for many years. They can come and tell my face (Kağıt üzerinde bir değerlendirme yapılmasına gerek yok. Birbirimizi çok iyi tanıyoruz çünkü uzun yıllardır birlikte çalışıyoruz. Gelip yüzüme söyleyebilirler.)”

4.1.2.3 Accuracy Problems

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the accuracy problems of the current teacher evaluation process in the school are given in Table 4.5.

Table 4.5

Codes for Accuracy Problems

<i>Theme 3. Accuracy</i>
<i>Misinterpretation of Teacher Effectiveness</i>
Competitive environment
Negative aspects of the observation process
<i>Bias Identification</i>
Biased judgements based on experience
Biased judgements based on personal relations
<i>Valid Judgements</i>
Not reflecting the complexity of the teaching/learning process
Ignoring other important domains of learning

Accuracy term is used to determine whether the information produced by the evaluation is profound enough to make decisions and make judgments. While ensuring accuracy, evaluation methodology should fit with the purpose of the evaluation and should be appropriate for the evaluates and the context in which they work (Howard & Gullickson, 2009). Problems with the accuracy of the evaluation are given under four subheadings: “Misinterpretation of Teacher Effectiveness,” “Bias Identification and Management,” and “Reliable Information.”

4.1.2.3.1 Misinterpretation of Teacher Effectiveness

Evaluations should promote valid judgements to minimize misinterpretation by using multiple data sources and linking the judgements to the purpose (Howard &

Gullickson, 2009). All the teachers stated that their effectiveness was misinterpreted; therefore, misinterpretation was causing problems for validity orientation, which has a serious effect on accuracy. Most of the teachers ($f=14$) pointed out that reason for the misinterpretation was the competitive environment; on the other hand, some teachers ($f=7$) pointed out the reason as behaving differently in the observation process. Accordingly, experts also considered the competitive environment ($f=5$) as the main reason for misinterpretation and teachers behaving differently in the observation process ($f=6$) as a problematic issue for validity orientation and consequently for accuracy. Principals ($f=6$) also stated that teachers behave differently than usual in the observation process, making the observer misinterpret the observation results.

A teacher (T1 *four years of experience*) explained the misinterpretation problem due to competitive environment as *"Everyone is compared to each other. Being successful is somewhat like the ego war between teachers. If you have achieved significant success in your class, this is the success of your class. Should not be compared to others. I don't find it right to make a decision about me by making comparisons with others. (Herkes birbiriyle kıyaslanıyor. Başarılı olmak öğretmenler arasında ego savaşı gibi. Sınıfınızda önemli bir başarı elde ettiyseniz, bu sınıfınızın başarısıdır. Başkalarıyla karşılaştırılmamalı. Benim hakkımda başkalarıyla kıyaslama yapılarak karar verilmesini doğru bulmuyorum)"* One expert explained the same problem as:

<i>Monitoring student development is essential thing in school. In other words, it is an indicator of how the teacher increases the student's success differently from the beginning. Therefore, it is not meaningful to compare teachers with other teachers by the average success of all students in a class; there are many students with different characteristics. (E6 five years of experience).</i>	<i>Bence öğrencinin gelişimini izlemek okuldaki en önemli şeydir. Yani aslında öğretmenin bir öğrencinin başarısını başlangıçtan farklı bir şekilde nasıl arttırdığı bir göstergedir aslında. Bu nedenle, öğretmenleri bir sınıftaki tüm öğrencilerin ortalama başarısı ile diğer öğretmenlerle karşılaştırmak anlamlı değil, farklı özelliklere sahip birçok öğrenci var sınıflarda. (E6 beş yıllık tecrübe)</i>

One teacher (T14 *thirteen years of experience*) reported misinterpretation problems due to behaving differently in the observation process, “*Although I think that an observation process is important, I may behave differently and not be objective if I know someone observing my behavior. (Gözlem sürecinin önemli olduğunu düşünmeme rağmen, birinin beni gözlemlediğini biliyorsam farklı davranabilir ve objektif olamayabilirim.)*” One expert (E6 *five years of experience*) explained the same problem as: “*During the observation, the teacher knows that her performance is being evaluated, and she will try to show her best performance. This situation causes the teacher to show or disregard many of her behavior different from the routine. (Gözlem sırasında, öğretmen performansının değerlendirildiğini bilir ve gözlemciye en iyi performansını göstermeye çalışır. Bu, öğretmenin rutinde yaptığı pek çok davranışı göz ardı etmesi ya da değiştirmesine neden olur.)*” Another teacher explained the negative aspects of the observation process as:

*Obviously, I don't find it very efficient for my principal to attend class. I think the teacher in the class behaves differently than typically. In a private school, anxiety is too high, and maybe things would be different if the process was different. We are human, and it feels uneasy about being observed. We constantly think that these observation results will be used against us. For that reason, perhaps we are trying to explain better or being more positive when we are reacting to the student. I mean, I don't think I'm showing my actual performance. (T10 *four years of experience*).*

*Açıkçası, müdürün derse katılmasını çok verimli bulmuyorum. Sınıftaki öğretmenin normalden farklı davrandığını düşünüyorum. Özel bir okulda kaygı çok yüksektir ve belki de işleyiş farklı olsaydı bir şeyler farklı olurdu. Biz insanız ve gözlemlenmek huzursuz hissettiriyor. Bu gözlem sonuçlarının bize karşı kullanılacağını düşünüyoruz sürekli olarak. Bu nedenle, öğrenciye tepki verirken belki de daha iyi açıklama yapmaya veya daha pozitif olmaya çalışıyoruz. Yani, ben gerçek performansımı gösterdiğimi sanmıyorum. (T10 *dört yıllık tecrübe*)*

4.1.2.3.2 Bias identification

Another problem that is affecting validity orientations and therefore affecting accuracy is defined as bias identification. Perception or beliefs held by the evaluator that influences the evaluator's judgment, which is not related to the teacher's performance, are called bias (Howard & Gullickson, 2009). Bias affects accuracy negatively because it undermines the fairness of the evaluation, distorts the data gathering

process, and corrupts decisions or actions. Most of the teachers ($f=14$) and experts ($f=4$) mentioned that teachers were evaluated differently according to their experiences and that evaluations were made in favor of experienced teachers, thus not free of bias. Teachers reported that teachers working for a long time in this school were not receiving much negative criticism because of the respect given to them by the principals. Experts also explained that experienced teachers' behaviors in this school were mostly ignored while novice teachers were constantly supervised. According to teachers and experts, this situation made novice teachers feel insecure and let the experienced teacher feel safe, comfortable, and resist change. One teacher explained this problem as:

[...] academic success is significant in this school, but I am an experienced teacher, and I can apply what I know without considering the average success of my class. Nobody warns me because I did this. However, the novice teacher is shaped according to this, and then the novice teacher is only teaching for academic success (T8 eighteen years of experience).

[...] akademik başarı bu okulda çok önemli, ama ben deneyimli bir öğretmenim ve bildiklerimi sınıfımın sınav başarı ortalamasını dikkate almadan uygulayabilirim. Bunu yaptım diye kimse beni uarmıyor. Ancak yeni öğretmen buna göre şekilleniyor ve bu kez yeni öğretmen sadece akademik başarı için ders işliyor (T8 on sekiz yıllık tecrübe).

Another teacher explained this problem in detail:

Most of the time, when I want to use the instructional methods or materials, I learned from the training I attend, the experienced teachers in my department say, "this will not work in this school; you are not experienced, you do not know." This time, when I insist and apply it in my classroom, I am faced with warnings from the principals that there is noise in the classroom. On the other hand, experienced teachers continue to practice what they knew twenty years ago, and I have never seen that they have received any warning. It feels like I must stop trying anything, but I haven't given up trying yet. (T19 three years of experience)

Çoğu zaman, katıldığım eğitimlerden öğrendiğim öğretimsel yöntemleri ya da materyalleri kullanmak istediğimde, zümremdeki tecrübeli öğretmenler "bu okulda işe yaramaz, deneyimsizsiniz tabi bilmiyorsunuz" diyorlar. Bu sefer sınıfımda ısrar edip uyguladığımda, sınıfta gürültü oluyor diye yöneticilerden gelen uyarılarla karşı karşıya kalıyorum. Öte yandan, deneyimli öğretmenler yirmi yıl önce bildiklerini uygulamaya devam ediyorlar ve daha hiç uyarı aldıklarını görmedim. Bu bana hiçbir şey denemeyi bırakmam gerektiğini hissettiriyor ama daha henüz denemekten vazgeçmiş değilim. (T19 üç yıllık tecrübe)

Another problem mentioned by teachers ($f=12$), principals ($f=5$), and experts ($f=4$) is making interpretations based on personal relations. The participants considered this problem also as biased identification. Teachers defined this problem as schools were also a place of socialization for teachers, and relationships were strengthening over time. Thus, teachers' fear of criticizing their friends stemmed from the fear of disrupting their relationship with their friends. On the contrary, the negative relations between the teachers causing biased evaluations were also stated as a problem by some of the teachers. Principals also stated that this problem occurred because the teachers were not open to criticism, particularly to criticism from her friend; personal relations were more important than professional relations in this school. Therefore, that criticism of her friend is perceived as a betrayal. Similarly, experts also explained this problem as principals did. An example excerpt is as follows: "*Teacher fears that relationship with another teacher-friend has deteriorated. In general, this results in each other closing down their gaps. No teacher at this school would tell you bad about another teacher even teachers who are head of departments (Öğretmen başka bir öğretmen arkadaşıyla ilişkisinin kötüleşmesinden korkuyor. Genel olarak bu birbirlerinin eksiklerinin üstünü örtmek olarak sonuçlanıyor. Bu okuldaki hiçbir öğretmen, zümre başkanı olan öğretmenler bile size bir öğretmen hakkında kötü bir şey söylemez)*" (E2 six years of experience). Similarly, one teacher explained this problem:

We are afraid to criticize the most crucial problem of integrated societies. In particular, we cannot criticize the people we have a personal relationship. We must get rid of the idea that "a teacher loved by everyone is a good teacher." At my old school, my supervisor observed the teachers very often, and I would not be disturbed as we were very used to the evaluation process, and the results were purely for improving us. I learned a lot to improve myself. If you do something negative in this school, you are being judged; I wish we were open to criticism. (T12 seven years of experience).

Entegre toplumların en önemli sorunu olan eleştirmekten korkuyoruz. Kişisel bir ilişki kurduğumuz insanları eleştiremiyoruz. "Herkes tarafından sevilen bir öğretmenin iyi bir öğretmen olduğu" fikrinden kurtulmalıyız. Eski okulumda amirim öğretmenleri sık gözlemlerdi ve değerlendirme sürecine çok alışık olduğumuz için hiç rahatsız olmazdım. Sonuçlar tamamen bizi geliştirmeye yönelikti. Hiç tedirginlik duymadım. Kendimi geliştirmek için çok şey öğrendim. Eğer bu okulda olumsuz bir şey yaparsan, yargılanıyorsun, keşke eleştiriye açık olsaydık. (T12 yedi yıllık tecrübe).

4.1.2.3.3 Valid Judgements

Another threat that may prevent accuracy is providing valid judgements. Validity refers to the degree to which judgments about a person's performance are reliable, and data can be obtained from more than one source to provide validity to the judgements. Based on the procedures of the evaluation system used by the evaluator, judgments in evaluation can be derived from multiple sources, including classroom observations, interviews, student surveys, peer reviews, portfolio reviews, project reviews, and student achievement data (Howard & Gullickson, 2009).

Almost all the teachers ($f=19$) considered using lesson plans as a single data source to understand the teacher's performance in the classroom as a problem. Because they thought that the lesson plans did not reflect the complexity of the teaching/learning process, they also stated that they did not receive any feedback regarding the quality of the lesson plans they prepared. Only one teacher (T6) said that she was satisfied with the evaluation of the lesson plans and explained that she had designed lesson plans in a very detailed manner; she researched before the lesson and tried very hard to comply with the plan. One teacher (T11_{twenty years of experience}) explained why using lesson plans is problematic *"The lesson plan is crucial, but the classrooms teacher is preparing plans for four lessons, and this is very difficult. It is hard to do right. If we cannot do it right, our plans are inadequate to reflect the process. I also think a plan on paper does not reflect my effort to teach in my classroom. (Ders planı çok önemlidir, ancak sınıf öğretmeni dört ders için planlar hazırlıyor ki bu bence çok zor. Yani bunu doğru yapmak zor. Eğer bunu doğru yapamazsak, planlarımız süreci yansıtmak için yetersiz kalıyor. Ayrıca kâğıt üzerinde yazılı bir planın sınıftaki öğretme çabamı yansıtmadığını düşünüyorum.)"*

One of the teachers also stated the problem as:

<i>I do not find it right to be evaluated only with a plan. You cannot always adhere to the daily plan because the class has a very active structure, and something else can happen at that moment. Anything can happen in the</i>	<i>Sadece planla değerlendirmeyi doğru bulmuyorum. Her zaman günlük plana bağlı kalamazsınız çünkü sınıfın çok dinamik bir yapısı var ve o anda başka bir şey yaşanabilir. Sınıfta her an her şey</i>

classroom at any time. When something is different from what you planned, you dream of something else, and you design it at that moment. Sometimes you stop what you have planned and leave it aside. (T14 thirteen years of experience).

olabilir. Planladığınızdan farklı bir şey olduğunda, başka bir şey hayal edersiniz ve o anda tasarlıyorsunuz. Bazen planladığınız şeyi durdurup bir kenara bırakırsınız. (T14 on üç yıllık tecrübe).

Most of the teachers ($f=15$) and experts ($f=6$) also stated that using students' average achievement scores as a single data source did not reflect the complexity of the instruction process in the classroom; therefore, this data source does not reflect the real performance of teachers. One of the teachers (T7 fifteen years of experience) explained this problem, "I do not think that collective success is my success. Performance is not simple enough to be measured by a test. I am successful with what I do in the classroom and what I contribute to the students. (Kolektif başarının benim başarıım olduğunu düşünmüyorum. Başarım bir test ile ölçülecek kadar basit değil. Ben sınıfta yaptıklarım ve öğrencilere kattıklarım ile başarılıyım.)" Similarly, another teacher used the following statements:

Teaching is not just about increasing academic success. Why should my effectiveness be measured by student success? For example, I had a lot of students who could not express themselves, and I have been trying for them for two years. Are we going to measure this effort with a school-wide exam? (T16 six years of experience)

Öğretim sadece akademik başarıyı arttırmak değil ki. Benim etkililiğim öğrenci başarısıyla neden ölçülsün? Örneğin kendini ifade edemeyen bir sürü öğrencim vardı onlar için iki yıldır çabalıyorum. Bu çabayı okul geneli sınavla mı ölçeceğiz. (T16 altı yıllık tecrübe)

Another teacher's explanations for a different application in the school are remarkable. She explained this process in detail as:

In our school, there are two groups of students. One group of students attends this school after achieving an exam, and other groups continue to primary school from kindergarten without taking any exam. I can compare the two groups, and the ones attending after passing an exam are more successful than the others.

Okulumuzda iki öğrenci grubu vardır. Bir grup öğrenci bir sınava girdikten sonra bu okula giriyor ve diğer grup anaokulundan herhangi bir sınava girmeden ilkokula devam ediyor. Bu iki grup arasında çok kolay da kıyaslama yaparım sınavı geçtikten sonra katılanlarla kesinlikle daha başarılılar. Bu sınavlı gruba bir

When I was teaching the students who passed the exam, I was an excellent teacher. Now I'm teaching students from our kindergarten. Although I've made three times more effort now, the students can't achieve the desired success. So now I am a bad teacher. (T12 seven years of experience).

şeyler öğretirken ben çok iyi bir öğretmendim. Şimdi anaokulumuzdan gelen öğrencilere ders veriyorum. Üç kat daha fazla çaba göstermeme rağmen, öğrenciler istenen başarıyı elde edemiyorlar. Şimdi ben kötü bir öğretmenim o halde. (T12 yedi yıllık tecrübe)

Another teacher expressed her discomfort with the application of school-wide exams as follows:

I think 20 questions on an exam do not show my class success. I learned from a master's degree that you must do piloting before the actual application of a test. It's not possible here. I do not believe that these exams are prepared to evaluate the real success of my students. The difficulty of these exam items is related to who prepares for the exam. The academic achievement of students should not represent 90% of teacher success. (T7 fifteen years of experience).

Bir sınavda sorulan 20 tane sorunun benim sınıftaki başarıyı göstermediğini düşünüyorum. Yüksek lisans derecesinden, bir testin gerçek uygulamasında bile önce pilotunun yapılması gerektiğini öğrendim. Burada bu mümkün değil. Bu sınavların öğrencilerimin gerçek başarısını değerlendirdiğine inanmıyorum. Bu sınav maddelerinin zorluğu, sınavı kimin hazırladığına bağlıdır. Öğrencilerin akademik başarısı öğretmen başarısının %90'ını temsil etmemelidir. (T7 on beş yıllık tecrübe)

Teachers (f=12) explained that using students' average achievement scores as a single data source ignores other important learning domains, such as the affective domain, which must be developed especially at early ages. Most of the experts (f=6) also stated this as a problem. They explained that in addition to academic achievement, it was important to support students' social and affective development and help them gain the skills they will need in their future lives. One of the teachers (T9 three years of experience) stated the problem, *"I don't think it's right to use student achievement as a single source. No one knows what I do for affective domain. Then it was up to my conscience to develop this domain (Öğrenci başarısını tek bir kaynak olarak kullanmanın doğru olduğunu düşünmüyorum. Duygusal gelişim için neler yaptığımı kimse bilmiyor. O zaman bu boyutu geliştirmek benim vicdanıma kaldı.)"* Another teacher explained the insufficient support of the other important learning domains as follows.

Being judged by the cognitive domain alone means we've lost some of the kids in our classrooms. According to the exam results, the administration has a general opinion about each teacher, and in fact, a teacher image has been formed. The teacher who brings more academic success is good. Why do we only follow academic achievements? Let's see if we support the emotional aspects of students. (T15 four years of experience)

Sadece bilişsel alana göre değerlendirilmek, sınıflarımızdaki bazı çocukları kaybettiğimiz anlamına gelir. Sınav sonuçlarına göre yönetimin her öğretmen hakkında genel bir görüşü var ve aslında bir öğretmen imajı oluşmuş durumda. Daha fazla akademik başarı getiren öğretmen iyi öğretmendir. Neden sadece akademik başarıları takip ediyoruz? Bakalım öğrencilerin duygusal yönlerini destekliyor muyuz? (T15 dört yıllık tecrübe)

4.1.3 Recommendations for Developing Effective Teacher Evaluation Model (Research Question 3)

The data to answer the third research question regarding recommendations for an effective teacher evaluation process was gathered through interviews conducted with teachers, principals, and experts working in that school. Participants were asked to propose suggestions to solve the problems they mentioned before. The relevant codes provided by the content analysis were organized under the headings “evaluation process” and “use of evaluation data.”

4.1.3.1 Recommendations for Evaluation Process

With the help of the answers received as a result of the interviews, those suggested to be in the process of the teacher evaluation model are grouped under four headings which are "purpose of evaluation," "evaluation schedule," "evaluator features," and "enabling evaluation dynamics."

4.1.3.1.1 Recommendations for the Purpose of Evaluation and Qualifications

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the purpose of the evaluation model and the qualification areas to be evaluated are given in Table 4.6.

Table 4.6*Codes for The Purpose of The Evaluation Model and The Qualification Areas*

<i>Theme 1. Evaluation Process</i>
<i>Purpose of Evaluation</i>
Evaluating and developing teacher qualifications
Improving student's achievement
<i>Teacher Qualification to be assessed</i>
Communication and collaboration
Instruction
Service to the school
Planning and preparation
Monitoring and managing learning
Professional development

One of the crucial problems of this case is the lack of a specific rationale for teacher evaluation. According to all the participants, the rationale of that teacher evaluation model should be both evaluating and developing teacher qualifications. Teachers, principals, and experts were also asked to specify what qualifications should be evaluated. As a result, they mentioned many qualifications and why they were essential to be evaluated. Most of the teachers ($f=16$), all the principals ($f=9$), and experts ($f=7$) stated that communication and collaboration were some of the most important qualifications that should have been evaluated and developed in this school. Some of the participants (T1, T4, T5, T9, T10, T19, P1, P3, P4, E1, E3) explained that they believed a teacher, who can communicate positively and effectively, can work in collaboration with other teachers and would also have been effective in the profession. Some participants (T2, T3, T6, T11, T20, P2, P4, P6, E5, E6) highlighted that teachers should have developed effective communication skills with their colleagues, students, and principals. A teacher explained the importance of communication and collaboration as follows:

Collaboration between teachers is very important because our common goal is to bring students to a good level. Everyone must work together and collaboratively, whether it is related to classes or extracurricular

Öğretmenler arasında işbirliği çok önemli çünkü hepimizin ortak amacı öğrencileri iyi bir seviyeye getirmek. Herkes gerek derslerle ilgili olsun gerekse ders dışı etkinlikler olsun bir arada ve işbirliği ile çalışmak

activities. An indispensable part of cooperation is to use communication effectively. That is, teachers need improvement and change in the field of communication and cooperation. (T10 four years of experience).

durumunda. İşbirliği yapmanın vazgeçilmez parçası da iletişimi etkili kullanmak. yani öğretmenlerin iletişim ve işbirliği alanında gelişmeye ve değişmeye ihtiyacı var. (T10 dört yıllık tecrübe).

One teacher explained this as follows.

There is a need for a teacher who can communicate and work harmoniously together. In this school, there are many teachers who graduated from university with first place and cannot communicate. But whether these people are good teachers or not, I think, is a controversial issue. I think being a good teacher means having high communication skills. Teachers should be in good communication with us and with students, parents, and administrators. (T2 five years of experience)

İletişim kurabilen ve birlikte uyumlu çalışabilen öğretmene ihtiyaç var. Bu okulda üniversiteden birincilikle mezun olmuşve iletişim kuramayan çok sayıda öğretmen var, ancak bu insanların iyi öğretmen olup olmadığı bence tartışmalı bir konu. Bence iyi öğretmen olmak yüksek iletişim becerilerine sahip olmak demektir. Öğretmenler sadece bizimle değil öğrenciler, velilerle ve idarecilerle de iyi iletişim halinde olmalıdır. (T2 beş yıllık tecrübe)

One principal (P6 sixteen years of experience) stated the importance of communication skill as “The teacher should not be seen only as increasing the academic achievement of the student. What it brings to the student in the affective dimension is very important, and this is achieved through communication in social life, so the teacher should always be in good communication with the student (Öğretmen sadece öğrencinin akademik başarısını arttıran kişi olarak görülmemelidir. Duygusal boyutta öğrenciye kazandırdıkları çok önemlidir ve bu sosyal hayatta iletişim yoluyla sağlanır, bu yüzden öğretmen her zaman öğrenci ile iyi iletişim içinde olmalıdır.)”

Most of the teachers ($f=15$), all the principals ($f=9$), and experts ($f=7$) conveyed that evaluating the effectiveness of instruction was one of the most important qualifications that should have been evaluated and developed in this school. Some participants (T1, T2, T4, T7, T9, T11, T13, T19, P2, P3, P4, P9, E1, E2, E3, E7) explained that it was not possible to evaluate the teacher if the instruction part was ignored entirely. Some of the teachers (T1, T2, T3, T5, T6, T16, T19, T20) also argued that if the effort of

classroom teachers in teaching young children was not observed, an objective assessment cannot be made. One teacher explained this as follows:

Sometimes there are times in the classroom when you must do something completely different and momentary instead of doing something you always planned. These children are very young, and young children need a lot of attention. We are trying in the classroom, but it is not known how much effort we have in the classroom. Therefore, observations are needed from inside the classroom, not outside the door. (T16 six years of experience)

Bazen sınıfta her zaman planladığınız bir şey yapmak yerine, tamamen farklı ve anlık bir şey uygulamak zorunda olduğunuz zamanlar vardır. Bu çocuklar çok küçük ve küçük çocuklar çok fazla ilgiye ihtiyaç duyuyor. Bizler sınıfta gerçekten çok çabalıyoruz ama sınıfta ne kadar çaba gösterdiğimiz bilinmiyor. Bu nedenle, kapının önünde değil, sınıfın içinden gözlemlere ihtiyaç var. (T16 altı yıllık tecrübe)

Another teacher also explained as follows:

Evaluating a teacher regardless of what is happening in the classroom means evaluating the result, not the process. When evaluating students, we always emphasize how important it is to evaluate the process. This situation is equally important to us. Sometimes, the general exam results in the school are considered equal to the success of the teacher. However, we do much with the students in the class, and sometimes it may not be reflected in the exam results. (T5 seventeen years of experience)

Bir öğretmeni sınıfta olup bitenlerden ayrı olarak değerlendirmek, süreci değil, sonucu değerlendirmek anlamına gelir. Öğrencileri değerlendirirken süreci değerlendirmenin ne kadar önemli olduğunu her zaman vurguluyoruz. Bu bizim için aynı derecede önemli. Bazen okul içindeki genel sınav sonuçları öğretmenin başarısına eşit kabul ediliyor. Ancak, sınıftaki öğrencilerle gerçekten çok şey yapıyoruz ve bazen bu sınav sonuçlarına yansımıyor. (T5 on yedi yıllık tecrübe)

Most of the teachers ($f=14$), principals ($f=7$), and experts ($f=5$) stated that service to the school was another critical qualification that should have been evaluated. According to participants, service to school consists of professional and in-school responsibilities of teachers such as making contributions to the development of the school, participating in the activities carried out in the school, keeping duties in the school, and taking responsibility in ceremonies. Some of the teachers (T1, T2, T4, T6, T9, T15, T16, T17, T19) explained that there should have been a difference between

the teacher who does these duties and responsibilities thoroughly and the teacher who does not. Furthermore, some principals (P2, P4, P8) stated that the school was very crowded, and to ensure the safety of the students, especially in the corridors, it was crucial for the teachers to perform their duties and responsibilities with whole motivation. The experts (E1, E2, E3, E4, E7), on the other hand, mentioned that these duties should have evaluated, and it was essential to distribute equally for everyone to fulfill their duties properly. One of the teachers explained as follows:

I always take care to fulfill my duties and responsibilities completely. It is vital to come to the school before the student and walk around the school without getting any harm. Similarly, attending ceremonies or being a supervisor in exams. I know that not everyone is as attentive as a few friends or me. Sanctions should be applied to these people. It may be used by me when I'm missing. In fact, I am sure that the managers make a lot of observations on this issue. The problem is that everyone knows very well who the teacher is who does not do her job but does nothing (T9 three years of experience)

Her zaman görev ve sorumluluklarımı eksiksiz yerine getirmeye özen gösteririm. Öğrenciden önce okula gelip onların okulun etrafında zarar görmeden dolaşmalarını sağlamak çok önemlidir. Aynı şekilde törenlere katılmak ya da sınavlarda gözetmen olmak. Herkesin ben veya birkaç arkadaşım kadar özenli olmadığını biliyorum. Bu kişilere yaptırımlar uygulanmalıdır. Eksik olduğum zaman bana da uygulansın. Aslında, yöneticilerin bu konuda çok fazla gözlem yaptıklarından eminim. Sorun şu ki herkes görevini yapmayan öğretmenin kim olduğunu çok iyi biliyor ama hiçbir şey yapmıyor (T9 üç yıllık tecrübe)

One of the principals stated:

It can be said that a teacher who fulfills her duties and responsibilities inside the school and outside her class is very successful and attentive in the classroom and other fields. There are many examples of this in our school. I think that fulfilling these responsibilities is directly proportional to being an effective teacher. All the teachers who are effective fulfill their duties at school. (P3 seventeen years of experience)

Okul içindeki yani kendi sınıfı dışındaki görev ve sorumluluklarını yerine getiren bir öğretmenin sınıfta ve diğer alanlarda çok başarılı ve özenli olduğunu söylenebilir. Bunun pek çok örneği var okulumuzda. Ben bu sorumlulukların yerine getirilmesinin etkili bir öğretmen olmakla doğru orantılı olduğunu düşünüyorum. İyi dediğim tüm öğretmenler okuldaki sorumluluklarını tam olarak yerine getiriyorlar. (P3 on yedi yıllık tecrübe)

Most of the teachers ($f=14$), all the principals ($f=9$), and experts ($f=7$) also mentioned that evaluating the planning and preparation process was another qualification to be evaluated. Among these participants, most of the teachers ($f=12$) noted that the evaluation of the lesson preparation process by only considering the lesson plans was insufficient in measuring the effectiveness of the instruction; instead, how well the preparation for the lesson can also be done by observing the teaching process. On the other hand, most of the experts (E1, E2, E3, E5, E7) and some principals (P3, P4, P6, P9) stated that the preparation of the course could be observed clearly with the help of lesson plans. One of the teachers explained as follows:

One of the most critical factors for the success of the course is the well-structured preparation phase. Before observing, we may be asked what kind of preparations or planning we are doing. However, I do not find it right just to examine the lesson plans and decide on the lesson process with the help of the lesson plan or see if the teaching fits the lesson plan. What happens in the classroom may be more complex and different from this lesson plan. In this case, not only my lesson plan but also the teaching itself should be observed. (T9 three years of experience)

Dersin başarısı için en önemli faktörlerden biri hazırlık aşamasının iyi yapılandırılmasıdır. Gözlemden önce, bize ne tür hazırlıklar ya da planlama yaptığımızı sorulabilir. Ancak, sadece ders planlarını incelemeyi ve ders planı yardımıyla ders sürecine karar vermeyi ya da sadece öğretimin ders planına uyup uymadığını gözlemlemeyi doğru bulmuyorum. Sınıf içinde yaşananlar bu ders planından daha karmaşık ve farklı olabilir. Bu durumda, sadece ders planım değil öğretimin kendisi de gözlemlenmelidir. (T9 üç yıllık tecrübe)

One of the principals stated:

Planning should be evaluated. The teacher sees preparing the lesson plan as a burden and taking the time to prepare the lesson plan is an unnecessary effort for them. However, if the teacher develops herself, a lesson plan should be designed. The classroom teacher should plan well what to teach that day and, if necessary, write plan b. (P9 six years of experience)

Planlama muhakkak değerlendirilmelidir. Öğretmen ders planı hazırlamayı kendine yük gibi görmektedir ve ders planı hazırlamaya zaman ayırmak onlar için gereksiz bir çaba. Ancak öğretmen eğer kendini geliştirecekse ders planı hazırlamalıdır. Sınıf öğretmeni o gün ne öğreteceğini o düzeyde çok iyi planlamalı ve gerekirse bu planını yazmalıdır. (P9 altı yıllık deneyim)

One of the experts indicated that:

Teachers learn to plan a lesson first while studying at university. This plan always has a format. The teacher does not understand that writing down the lesson plan only on paper is synonymous with having good planning skills. They need to rehearse what they wrote in the lesson plan. It means presenting all kinds of visuals, texts, tools, and materials to the student correctly. This mentioned qualification is an essential teacher qualification. The lesson plan is the assistant of the teacher. If I were a teacher, I would probably feel incomplete without my plan (E1 eight years of experience).

Öğretmenler, üniversitede okurken önce bir ders planlamayı öğrenirler. Hatta bu planlamanın hep bir formatı vardır. Bence öğretmenin anlamadığı şey aslında ders planını sadece kâğıt üstünde yazıp bırakmanın iyi planlama becerisine sahip olmakla eş anlam ifade ettiğidir. Ders planına yazdıkları şeyleri prova etmeleri gerekir. Öğrenciye sunulacak her türlü görseli, metni, araç ve materyali planlayarak doğru sırayla sunmak demektir. Bu çok önemli bir öğretmen yeterliliğidir. Ders planı öğretmenin asistanıdır. Ben öğretmen olsaydım muhtemelen planım olmadan eksik hissederdim (E1 sekiz yıllık tecrübe).

Half of the teachers (T2, T4, T6, T7, T10, T14, T16, T17, T18, T20) mentioned that monitoring and managing learning was also essential to evaluate and it was critical to monitor the development of students to take precautions. The other half of the teachers (T1, T3, T5, T8, T9, T11, T12, T13, T15, T19) mentioned that the general exams held in this school were not sufficient to measure student success and that the weaknesses of the students could not be determined correctly. In addition, some of the teachers (T4, T7, T10, T17, T20) and a few principals (P7, P9) argued that it would have been more accurate to follow the development of cognitive achievement of students at this age level as well as the development of the affective domain. On the other hand, all the principals ($f=9$) and most of the experts ($f=7$) mentioned that monitoring and managing the student's learning was an indispensable qualification to be evaluated. A teacher who thought that this qualification should not be evaluated due to the nature of the general exams explained her ideas as follows:

The monitoring and elimination of the deficiencies of the students are followed here only with the general exams. Like whether the student received a low in the previous exam

Öğrencilerin eksikliklerinin izlenmesi ve giderilmesi burada sadece genel sınavlarla takip ediliyor. Öğrenci bir önceki sınavda düşük mü aldı diğerinde yükseltti mi gibi. Bu okulda

or increased it in the other. It is doubtful whether the questions in the general exams held in this school are suitable for the cognitive level of the student, and those who prepare the questions do not consult us anyway. Many teachers do not trust these exams for this reason. (T15 *four years of experience*)

yapılan genel sınavlardaki soruların öğrencinin bilişsel düzeyine uygunluğu şüpheli ve soruları hazırlayanlar bize zaten danışmıyorlar. Pek çok öğretmen bu sınavlara bu nedenle güvenmiyor. (T15 *dört yıllık deneyim*)

Another who thinks that monitoring and managing the learning is essential to evaluate explained her ideas as follows:

In this school, the only exam that we can see the student's level of success is the general exams. Moreover, a teacher can observe the student's level of success in classroom practice. For example, I know that my student is missing both the general exam and in the classroom. Students' weaknesses are immediately noticed. A good teacher knows very well what the student in the class is lacking and tries hard to overcome these shortcomings. (T18 *ten years of experience*)

Bu okulda öğrencinin okulumuzdaki başarı düzeyini görebildiğimiz tek sınav genel sınavlardır. Dahası, bir öğretmen öğrencisinin sınıf uygulamalarındaki başarı seviyesini de gözlemleyebilir. Ben mesela hem genel sınavda hem de sınıfta öğrencimin eksik olduğunu biliyorum. Öğrencilerin zayıf yönleri hemen fark edilir. İyi bir öğretmen sınıfındaki öğrencinin hangi konuda eksik olduğunu çok iyi bilir ve bu eksikliği gidermek için çok çaba sarf eder. (T18 *on yıllık tecrübe*)

Professional development was also crucial for most principals ($f=7$), and all the experts ($f=7$) mentioned it as a qualification to be evaluated. In addition, most of the teachers ($f=13$) said that measuring this qualification was essential but may reveal some undesirable problems. For example, some teachers (T5, T8, T13, T14) defined the problem as feeling obliged to participate without a need for it; some other teachers (T1, T4, T11, T19) mentioned the issue as creating a competitive environment among teachers. Few of them (T9, T15, T18) defined the problem as some teachers may only attend these pieces of training just to provide their existence even though she does not take an active role in the training. On the other hand, some other teachers (T5, T9, T13, T17, T18) mentioned the difficulty of evaluating this qualification objectively. One of the teachers explained as follows:

Professional development should be evaluated, I think, as an essential qualification. No matter how many years of experience we have, the generation that comes each year is very different. Personally, I can only read the latest education news, and I have not attended the training any time soon, but I should. Participation in training is critical, and I think this is an important criterion to be evaluated. The teacher should not feel obligated and believe that it is necessary if they do it because they are compulsory, they cannot see the benefit. (T8 eighteen years of experience)

Mesleki gelişim kesinlikle değerlendirilmeli bence çok önemli bir yeterlik. Kaç yıllık tecrübemiz olursa olsun her yıl gelen nesil birbirinden çok farklı. Ben şahsen sadece eğitim ile ilgili en son haberleri okuyabiliyorum ve yakın zamanda eğitimlere katılmadım, ama katılmalıydım. Eğitimlere katılım çok önemli ve bence bu değerlendirilecek önemli bir ölçüt. Sadece öğretmen buna zorunlu hissetmemeli ve gerekli olduğuna inanmalıdır, eğer zorunlu olduğu için yaparlarsa faydasını göremezler (T8 on sekiz yıllık tecrübe)

One teacher expressed the difficulty of evaluating professional development as follows:

In other schools where professional development is measured, teachers attend training even though they do not develop themselves in this area. I know it from my old school. After just going to a seminar and watching the first sessions, I got to know many teachers leaving the seminar. In this case, professional development deviates from its true purpose. (T9 three years of experience)

Mesleki gelişimin ölçüldüğü diğer okullarda, öğretmenler kendilerini bu alanda geliştirmeye gitmedikleri halde eğitimlere katıldıklarını söylüyorlar. Ben eski okulumdan biliyorum bunu. Sadece bir seminere gidip ilk oturumları izledikten sonra seminerden ayrılan bir sürü öğretmen tanıdım. Bu durumda mesleki gelişim gerçek amacından sapıyor. (T9 üç yıllık tecrübe)

As mentioned before, most of the participants agreed that the teachers were being evaluated to improve student achievement. On the other hand, teachers and experts indicated that using students' average achievement scores as a single data source or evaluating teachers causes some problems like not evaluating objectively, neglecting the teaching process, and focusing only on the results. When participants asked for recommendations from some of the teachers ($f=9$), most of the principals ($f=5$) and experts ($f=4$) also mentioned that improving students' achievement was an important qualification to be evaluated. Some participants (T4, T5, T9, T13, T17, P2, P3, P4, P8, E1, E2, E4, E6) explained that teacher qualifications in improving students'

achievement could be evaluated when student achievement was not seen as a single source, it was essential where individual development of students could be monitored rather than an average achievement and in situations where the students were educated not just for academic success but within a holistic perspective. One of the teachers explained as

"We are teachers; of course, we do our best to increase the success of the student. But rather than the success of the whole class, students' individual achievement should be looked at to understand what we have contributed to this student individually. This situation should be considered both in academic and other fields. I don't find it right to evaluate collective success as my success (T5 seventeen years of experience)."

Biz öğretmeniz, elbette, öğrencinin başarısını artırmak için elimizden geleni yapıyoruz. Ancak sınıfın başarısı yerine öğrencilerin bireysel başarısına bakılmalıdır. Bu öğrenciye bireysel olarak ne tür katkılarda bulunduk. Bu hem akademik hem de diğer alanlarda böyle düşünülmelidir. Genel başarıyı başarımla değerlendirilmeyi doğru bulmuyorum. (T5 on yedi yıllık tecrübe)

The statement made by one of the teachers was remarkable:

They have to separate us from other branch teachers. The branch teacher does not spend as much time as we do, and we class teachers know every student very well. I think the exams do not see the student's strengths and weaknesses as well as we do. Problems arise when student test results are considered equivalent to teacher success. Because we, classroom teachers, are more than students' joint success. But of course, if an excellent student in our class had a significant drop in exams, this is a problem for the teacher. (T9 three years of experience).

Bizi diğer branş öğretmenlerinden ayırmaları lazım. Branş öğretmeni bizim kadar vakit geçirmiyor öğrenciyle ve biz sınıf öğretmenleri her öğrenciyi çok iyi tanıyoruz. Sınavların öğrencinin güçlü ve zayıf yönlerini bizim kadar iyi bilmediğini düşünüyorum. Öğrenci test sonuçları öğretmen başarısına eşdeğer kabul edildiğinde sorunlar ortaya çıkar. Çünkü biz sınıf öğretmenleri öğrencilerin ortak başarısından daha fazlasıyız. Ancak tabi Sınıfımızdaki çok iyi bir öğrenci sınavlarda büyük bir düşüş yaşıyorsa, bu öğretmen için bir sorundur. (T9 üç yıllık tecrübe).

One of the experts explained why it is not possible to determine individual student success with the help of general exams at school as follows:

Students' general exam success informs us about the gains that students cannot learn. The teacher

Öğrencilerin genel sınav başarısı bize öğrenciler tarafından öğrenilemeyen kazanımlar hakkında

tries to compensate for missing learning. In these exams, deficiencies are generally determined rather than individual success. The aim is not to determine the success of the teacher. A one-time measurement with the help of a single question for success does not actually provide enough information about the student's condition. Therefore, our general exams do not provide enough data about the student's individual achievement. We must define what we expect the students to acquire at each grade level as the minimum skills. If we measure and report these skills through exams, observations, or assignments in the process, then we can say that we follow the student's individual development at this school. (E6 five years of experience)

bilgi verir. Öğretmen eksik öğrenmeleri telafi etmeye çalışır. Bu sınavlarda eksiklikler bireysel başarıdan ziyade genel olarak belirlenir. Zaten amaç, öğretmenin başarısını belirlemek değildir. Bir başarı için tek bir soru yardımıyla bir kerelik bir ölçüm, öğrencinin durumu hakkında yeterli bilgi sağlamaz aslında. Bu nedenle yaptığımız genel sınavlar öğrencinin bireysel başarısı hakkında yeterli veri vermez. Öğrencilerin her sınıf düzeyinde kazanmasını beklediğimiz yani asgari becerilerin ne olduğunu tanımlamalıyız. Bu becerileri süreçteki sınavlar, gözlemler veya ödevler aracılığıyla ölçüp raporlarsak işte o zaman öğrencinin bu okuldaki bireysel gelişimini takip ediyoruz diyebiliriz. (E6 beş yıllık tecrübe)

4.1.3.1.2 Recommendations for Evaluation Schedule

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the schedule of the evaluation model are given in Table 4.7.

Table 4.7

Codes for The Evaluation Schedule

<i>Theme 1. Evaluation Process</i>
<i>Evaluation Schedule</i>
Not at the beginning of each academic period
One or two observations for each month
Three classroom observations for each semester

Most of the teachers ($f=14$) stated that they were busy adapting students to school at the beginning of each academic period; therefore, it would have not been correct to

make observations in the first months or weeks of each academic period. Moreover, the majority of the participants stated that these observations should have not been limited to a single observation, and it would have been accurate to observe different lessons of a teacher. Most of the teachers ($f=14$), principals ($f=7$), and experts ($f=6$) stated that it would have been uncomfortable for the teacher to make three observations in a month and making one or two observations every month would prevent the continuous division of courses.

4.1.3.1.3. Recommendations for Evaluator Features

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the evaluators' features are given in Table 4.8

Table 4.8

Codes for The Evaluator Features

<i>Theme 1. Evaluation Process</i>
<i>Evaluator Features</i>
Objective/Fair
Positive communication skills
Has teaching experience
Empathy skills
Assessment of knowledge and skills
Enough knowledge about school
Effective communication skill

A significant number of teachers ($f=19$) stated that they would have felt more comfortable in this process when they believe that an evaluator is objective and will not need to behave differently than they do in their classrooms during a classroom observation. Some teachers (T5, T9, T11, T13, T14, T19) also stated that the objectivity of the evaluator was critical to trusting the evaluation results. For example, one of the teachers explained:

Perhaps the essential feature of an observer is objectivity, the most important and possibly the most difficult. The person may not make an *Bir gözlemcinin belki de en önemli özelliği objektif olmasıdır. En önemlisi ve belki de en zoru. Kişi tamamen tarafsız bir değerlendirme*

utterly neutral assessment, but this is the only way I can trust someone's observation. I think the classroom environment will reflect the truth if the evaluator can be objective because the observer will also allow the teacher to be herself. Actually, there is a mutual situation here. (T11 twenty years of experience)

yapamayabilir, ancak benim bir kişinin gözlemine güvenebilmemin tek yolu bu. Değerlendirici objektif olabiliyorsa, sınıf ortamının da gerçeği yansıtacağını düşünüyorum açıkçası çünkü gözlemci öğretmenin kendisi olmasına da izin verecektir. Aslında burada karşılıklı bir durum var. (T11 yirmi yıllık tecrübe)

The majority of the teachers ($f=16$), more than half of the principals ($f=5$), and experts ($f=5$) stated it was essential that the evaluator is skilled in terms of communication. Teachers and principals defined these skills as "positive communication skills"; on the other hand, experts defined these skills as "effective communication skills." Teachers and principals stated that teachers would have felt more comfortable expressing themselves when the evaluators can communicate positively. Some of them (T1, T3, T4, T5, T10, T11, T12, T16, T18, T20, P1, P4, P7) explained that this would enable them to approach the evaluations positively and that it was essential to establish good communication with the evaluator to support their development. For example, one of the teachers pointed out that:

Positive communication is also crucial in our school. We are a group of teachers who can understand each other because we communicate well. We use positive language to criticize each other, and we know that this criticism is for our good. The evaluator must establish positive communication to create an environment where I can internalize his criticism of me. For example, if someone from my department evaluates me, I feel comfortable. (T5 seventeen years of experience)

Olumlu iletişim de kurumumuzda çok önemlidir. İyi iletişim kurabildiğimiz için, birbirimizi iyi anlayabilen bir grup öğretmeniz. Olumlu dil kullanarak birbirimizi eleştiriyoruz ve bu eleştirinin kendi iyiliğimiz için olduğunu biliyoruz. Değerlendiricinin de bana olan eleştirilerini içselleştirebileceğim ortam yaratabilmesi için olumlu iletişim kurması gerekir. Ben mesela zümremden birileri beni değerlendirse kendimi rahat hissederim. (T5 on yedi yıllık tecrübe)

In addition, one of the teachers pointed out that:

The evaluator should be positive. Even her facial expression and mimics are essential. She must not be sullen and should not make you

Değerlendirici olumlu olmalıdır. Yüz ifadesi, mimikleri bile önemli. Asık suratlı olmamalıdır ve gözlemlendiğinizi hissettirmemelidir.

feel observed. Maybe she should join the class. I must know that you came to support me, not criticize me. The first observations can be tricky, but it may be more comfortable if I know she is there to help me. (T4 three years of experience)

Belki de sınıfa katılmalıdır. Beni eleştirmeye değil, beni desteklemeye geldiğini bilmeliyim. İlk gözlemler zor olabilir, ancak beni desteklemek için orada olduğunu bilirim daha rahat olurum. (T4 üç yıllık tecrübe)

Experts (E1, E3, E7) explained that evaluators should have been more effective communication skills than positive communication skills. They added as positive communication was not an obligation, but it was essential to communicate effectively to convey the information clearly. Some of the experts (E1, E4, E6) stated that at the beginning of effective communication, positive aspects could be emphasized but still, it was essential to convey the seriousness of the deficiencies resulting from the evaluation. One of the experts explained as:

I think we are not open to criticism as Turkish society. Especially teachers often misunderstand what we say to them. In fact, we should be able to tell the shortcomings we see directly as experts. In doing so, we must convey the message we want to give correctly, but we should not soften the content of the message. When you try to be positive, the teacher does not understand you or misunderstands and criticizes you. I think communication should be effective, not positive. (E7 three years of experience)

Bence bizler Türk toplumu olarak eleştiriye açık değiliz. Özellikle öğretmenler onlara söylediklerimizi çoğu zaman yanlış anlayabiliyorlar. Aslında doğrudan uzman olarak gördüğümüz eksiklikleri söyleyebilmeliyiz. Bunu yaparken, doğru bir şekilde iletmek istediğimiz mesajı iletmeliyiz, ancak mesajın içeriğini yumuşatmamalıyız. Olumlu olmaya çalıştığınızda, öğretmen sizi anlamaz ya da yanlış anlayıp sizi eleştirir. Bence iletişim olumlu değil etkili olmalı. (E7 üç yıllık tecrübe)

Most teachers (f=13) indicated that evaluators should have been experience in teaching. Some of the teachers (T1, T5, T10, T15, T17, T20) explained that a person who observes themselves in the classroom must have known the teaching profession to understand the classroom environment and make more accurate evaluations. Some teachers (T2, T3, T9, T17, T19, T20) explained that they would have been feeling more comfortable, especially during the observation, if they knew that the evaluators were experienced in teaching primary schools. A teacher explained the importance of the evaluators having teaching experience as follows:

[...] for example, if a person does not know the developmental characteristics of a primary school student, she may find the current situation in the classroom as very complex and erratic. We have encountered such criticism in the past. When a person who has never had a classroom teaching experience passed in front of the classroom, she complained that there was a lot of noise coming from the classroom, there was no teaching in this classroom, and worst of all, her complaint was taken seriously. For this reason, it is crucial to have a person who knows primary school students and even works as a primary school teacher. (T17 fifteen years of experience)

[...] örneğin, bir kişi bir ilkökul öğrencisinin gelişimsel özelliklerini bilmiyorsa, sınıftaki o anki durumu çok karmaşık ve düzensiz bulabilir. Geçmişte bu tür eleştirilerle karşılaştık. Daha önce hiç sınıf öğretmeni olmayan bir kişinin sınıfın önünden geçerken, sınıftan çok fazla gürültü geliyor diye bu sınıfta öğretim yapılmıyor diye şikâyet etti ve en kötüsü, şikâyeti ciddiye alındı. Bu nedenle, ilkökul öğrencilerini tanıyan ve hatta ilkökul öğretmeni olarak çalışan bir kişinin olması çok önemlidir. (T17 on beş yıllık tecrübe)

More than half of the teachers (f=13) mentioned that establishing empathy was a skill that must be possessed while making evaluations. Some of them (T1, T2, T5, T9, T11, T12, T16, T17, T20) explained the observer should have evaluated the instruction according to the classroom situation they are in, and this may be possible by people with high empathy skills. For example, one of the teachers explained:

For example, I think an evaluator is a very knowledgeable and excellent teacher, but more importantly, she lacks empathy and positive communication. I would not want to be evaluated by such a person. If she cannot put herself in my place, she cannot correctly assess what I am doing in the classroom, and it incorrectly determines where I am missing. (T1 four years of experience)

Şimdi, örneğin, çok bilgili ve çok iyi bir öğretmen olan, ancak daha da önemlisi, empati ve olumlu iletişim eksikliği olan bir değerlendirici olduğumu düşünüyorum. Böyle bir kişi tarafından değerlendirilmek istemem. Kendini benim yerime koyamazsa, sınıfta yaptığım şeyi doğru bir şekilde değerlendiremez ve dahası, nerede eksik olduğumu yanlış bir şekilde belirler. (T1 dört yıllık tecrübe)

As teachers (f=10), principals (f=6), and experts (f=6) also agree that an evaluator should have had the qualifications and skills required, such as using the tools effectively, fulfilling the requirements of the observer role, and making accurate evaluations. Some participants (T4, T9, T10, T12, P1, P3, P9, E1) pointed out that the experienced evaluators, such as those having experience making observations or interviews by using observation or interview forms, could have conducted the

evaluation process more accurately. Some participants (T8, T14, T15, T17, P5, P8, E4, E5) stated that evaluators would have acquired these competencies with the necessary training and practical applications. One of the teachers gave the following explanation about the competencies that the evaluator should have.

The more experienced the evaluator is, the better she will master the process. People are doing postgraduate education here, and they have made observations or interviews before. I think these people dominate the evaluation process. A person who is already observing for the first time cannot immediately understand what to observe. Whoever has observed before knows very well what to look for in the class. (T10 four years of experience)

Değerlendirici ne kadar deneyimli olursa, sürece o kadar hâkim olacaktır. Burada lisansüstü eğitim yapan insanlar var ve kesinlikle daha önce gözlem veya görüşme yapmışlardır. Bu insanların değerlendirme sürecine hâkim olduklarını düşünüyorum. Zaten ilk kez gözlem yapan bir kişi neyi gözlemleyeceğini hemen anlayamaz. Kim daha önce gözlemlemişse neye bakacağını çok iyi bilir. (T10 dört yıllık tecrübe)

More than half of the teachers ($f=11$) and principals ($f=6$) stated that the evaluator should have had enough information about the school. The participants noted that evaluators should have been worked in the school and know the culture or system of the school. They also added that an external evaluator could not correctly evaluate because they do not recognize the school. Some of the participants (T3, T5, T7, T8, T11, T18, T19, P1, P6, P9) expressed the following problems that might arise from being evaluated by an external auditor: being felt under control, being uneasy, not being able to express themselves as they wish, and having difficulty in getting used to someone outside of the school. Few teachers (T12, T15, T20) have stated that the involvement of an external supervisor in a model that supports professional development would have been incomplete in determining the professional needs of teachers working in the school. One of the teachers (T20 *nine years of experience*) explained as “*A person outside the school could not make accurate evaluations more than the evaluations made by the people who knew the school (Okul dışındaki bir kişi, okulu bilen kişilerin yaptığı değerlendirmelerden daha doğru değerlendirme yapamaz.)*”

4.1.3.1.4 Recommendations for Enabling Evaluation Dynamics

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the evaluation dynamics are given in Table 4.9.

Table 4.9

Codes for The Evaluation Dynamics

<i>Theme 1. Evaluation Process</i>
<i>Enabling Evaluation Dynamics</i>
Training schedule for evaluators
Information meetings for the teacher being evaluated

Most of the teachers ($f=17$), principals ($f=9$), and experts ($f=7$) stated that the evaluators should have participated in training before the implementation of the evaluation model. They said that the evaluators should have received training to know how to use the data collection tools and what they should pay attention to during the evaluation process. For example, one of the teachers (T15 *seventeen years of experience*) pointed out, "Evaluators may not have much information and skills about the evaluation, but they can be competent in this field with the training they received. I do not want to be evaluated by a person who has not received such training. (Değerlendiriciler ilk başta değerlendirme hakkında çok fazla bilgi ve beceriye sahip olmayabilir, ancak aldıkları eğitimlerle bu alanda yetkin olabilirler. Şahsen, eğitim almamış bir kişi tarafından değerlendirilmek istemem.)"

Most of the teachers ($f=15$), principals ($f=7$), and experts ($f=5$) stated that the teachers who will be evaluated should have participated in information meetings before the implementation of an evaluation model. They mentioned that informing teachers about the evaluation model, especially its purpose was important for the effectiveness of applying the model. Another the teacher pointed out:

In this model, they must tell us what to do and how to do it. I mean, I do not have to learn how to observe with my effort, for example, by asking another friend. I need to know what happens from start to end of the observation.

Bu modelde bize ne yapacağımızın ve nasıl yapacağımızın söylemesi gerekiyor. Yani ben gözlem nasıl yapılır diye kendi çabamla örneğin bir başka arkadaşıma sorarak öğrenmemeliyim. Gözlemde baştan

There should be training at the very least, including the purpose of this evaluation model, data collection tools to be used, and even observation. (T16 six years of experience)

sona ne olacağını bilmem gerekiyor. En azından bu değerlendirme modelinin amacı, kullanılacak veri toplama araçları ve hatta gözlem hakkında bir eğitim olmalıdır. (T16 altı yıllık tecrübe)

4.1.3.1.5 Recommendations for the Evaluator

The codes of the recommendations that emerged as a result of the analysis of the opinions of teachers, principals, and experts about who the evaluator should be are given in Table 4.10

Table 4.10

Codes for The Evaluator

<i>Theme 1. Evaluation Process</i>
<i>Evaluator</i>
Principals
Classroom teacher
Teacher herself
Math/science teachers
School Experts
Students
Parents

Almost all the teachers ($f=19$), all the principals ($f=9$), and experts ($f=7$) stated that principals could have taken an active role in the evaluation process. Some teachers (T1, T2, T5, T7, T9, T11, T14, T18, T20) stated that school principals or assistant principals could collect data through observation and evaluate teachers since they have teaching experience. Some other teachers (T1, T3, T4, T7, T9, T12, T17, T19) expressed that if the school principal and assistant principals conducted the observation process, they would realize how much effort the teachers had made in the classroom and would provide them with more objective information about themselves. For example, one of the teachers stated:

I really want the principal to observe my class. I have known our principal from the moment she became a teacher,

Okul müdürünün sınıfımı gözlemlemesini gerçekten istiyorum. Müdürümüzü öğretmen olduğu

and during her observation, I found her approach in the classroom very correct. What she told us about her teaching experience and positive communication in the classroom made me feel comfortable during the observation process. Assistant principals can also make observations. The assistant principal has not made any observations yet, but he is also a teacher, and it is essential for me to get his feedback. (T11 twenty years of experience)

andan itibaren biliyorum tanıyorum ve gözlem sırasında sınıftaki yaklaşımını çok doğru buluyorum. Bize öğretmenlik deneyimi ve sınıftaki olumlu iletişimi hakkında anlattığı şeyler, gözlem sürecinde beni rahatlattı. Müdür yardımcıları da gözlem yapabilirler. Müdür yardımcısı henüz gözlem yapmadı, ama o da aynı zamanda bir öğretmen ve onun geri bildirimlerini almak da benim için çok önemli. (T11 yirmi yıllık tecrübe)

While some of the principals (P1, P3, P4, P5) considered themselves competent in evaluating teachers, some of the experts (E1, E2, E4, E5, E6) stated that principals should receive the necessary training and then they could observe teachers in the classroom effectively. One of the experts pointed out that:

[...] the principals here have been teachers and have elementary teaching experience, but evaluation requires other qualifications. These qualifications are not something that can be achieved so quickly. The person must master each item written in the observation or interview form and know the exemplary behavior reflected by these items. This is only possible with long-term trainings. (E5 two years of experience)

[...] buradaki yöneticiler daha önce öğretmenlik yapmış ve ilköğretim öğretmenliği deneyimine sahip, ama değerlendirme başka yeterlilikler de gerektiriyor. Bu yeterlilikler bu kadar çabuk kazanılabilecek bir şey değil. Kişi, gözlem veya görüşme formunda yazılı her bir şeye hâkim olmalı ve bu öğelerin yansıttığı örnek davranışları bilmelidir. Bu ancak uzun süreli eğitimlerle mümkündür. (E5 iki yıllık tecrübe)

Most of the teachers ($f=16$), principals ($f=8$), and experts ($f=6$) stated that other classroom teachers could collect data for the teacher evaluation process. Teachers indicated that it was essential for another teacher to observe her classroom instruction to make them aware of the weaknesses they did not realize before. On the other hand, they stated that it would have been an excellent opportunity to observe different teachers, which would lead them in their professional development. Experts and principals explained that teachers would learn a lot from each other's lesson observation, which was also essential in achieving teaching unity. Some teachers (T3, T5, T11, T14, T16, T17, T19) stated that if they were to be evaluated by a teacher

friend, they should have believed the objectivity of the teacher first. They also added that teachers who could not get along with each other might reflect on their personal problems during the observation process. Therefore, they should not observe each other's class.

Many teachers ($f=14$) stated that they could provide data on the quality of teaching by making self-assessment. Some principals (P2, P6, P8, P9) and experts (E1, E3, E4, E5, E7) also stated that they could evaluate themselves if the teachers received the necessary training. Principals and experts, who disagreed with this opinion, have generally said that they don't believe teachers can evaluate their own objectively. A teacher expressed her thoughts about self-assessment as follows:

<i>[...] perhaps the most important thing is that the teacher can evaluate herself. Because a teacher knows herself very well. If she knows how to criticize herself, there is no one who can better evaluate her teaching. Perhaps the best way to not be afraid of someone's criticism about you is to evaluate yourself first. (T6 six years of experience)</i>	<i>[...] belki de en önemlisi öğretmenin kendini değerlendirebilmesidir. Çünkü bir öğretmen kendini çok iyi tanır. Kendisini nasıl eleştireceğini bilirse, öğretmenliğini daha iyi değerlendirebilecek kimse yok. Belki de bir başkasının sizinle ilgili eleştirilerinden korkmamanın en iyi yolu önce kendinizi değerlendirmektir. (T6 altı yıllık tecrübe)</i>

The 3rd and 4th grade teachers ($f=8$) and teachers who will teach at these grade levels in the following years ($f=7$) stated that math and science teachers would make them aware of their competencies in science and mathematics courses by observing them, and they could correct them if they had misconceptions. The teachers (T2, T3, T5, T6, T8, T9, T11, T12, T16, T20) explained that it was crucial that teachers from different branches should support only in terms of field knowledge while observing their courses and that it would not be right to give opinions about classroom management. One of the teachers explained as:

<i>Teachers from different branches can say something about the field. It is helpful to be aware of new information in this field, especially in science and mathematics lessons, or to give me feedback if I am wrong or</i>	<i>Farklı branşlardan öğretmenler alan hakkında bir şeyler söyleyebilirler. Özellikle fen ve matematik derslerinde bu alandaki yeni bilgilerden haberdar olmamda veya yanlışım ya da eksikim varsa bana</i>

missing. For example, last year, I was teaching fourth-grade students and explaining the states of matter. I called the students solid, liquid, gas, and the kids told me about the fifth-sixth state. I went home and researched from different sites. They are more competent than me in the field, and I think they will contribute significantly to my teaching. (T7 fifteen years of experience)

dönüt vermede faydalı olur. Örneğin, geçen yıl dördüncü sınıf öğrencilerine ders veriyordum ve maddenin hallerini açıklıyordum. Öğrencilere katı, sıvı, gaz dedim ve çocuklar bana beşinci altıncı halden bahsettiler. Eve gittim ve farklı sitelerden araştırma yaptım. Alan açısından benden daha yetkinler ve bence öğretmenliğime büyük katkı sağlayacaklar. (T7 on beş yıllık tecrübe)

One of the teachers explained how to be evaluated by teachers from different branches as follows:

Teachers from different branches can evaluate. I think I have a lot to learn from these teachers. For example, in science and math, they can be beneficial in terms of field knowledge, but I do not let them criticize me for my classroom management skills in the classroom. Sometimes I ask my mathematician friends what a term means. Therefore, I think the opinions of teachers from different branches are significant. (T20 nine years of experience)

Farklı branşlardan öğretmenler değerlendirebilir. Bu öğretmenlerden öğrenecek çok şeyim olduğunu düşünüyorum. Örneğin fen ve matematikte, alan bilgisi açısından çok yardımcı olabilirler, ancak kendimi sınıftaki sınıf yönetim becerilerim açısından eleştirtmem. Bazen matematikçi arkadaşlarıma bir terim ne anlama geliyor? diye sorarım. Bu nedenle, farklı branşlardan öğretmenlerin görüşleri bence çok önemli. (T20 dokuz yıllık tecrübe)

Few of the teachers (T2, T6, T10, T15) stated that an expert working at the school could also evaluate teachers, while most experts ($f=5$) indicated that they were qualified to evaluate the teachers. Most of the teachers did not accept experts as evaluators because the experts had not graduated from classroom teaching, lack of teaching experience, and lack of positive communication.

Few teachers (T8, T13, T15) and principals (P2, P6, P9) stated that especially fourth-grade students could evaluate the teachers. Most teachers and principals and all experts noted that it was not correct for teachers to be evaluated by students. While supporting the view that students should not be evaluators, it was pointed out that especially primary school students could not make objective evaluations due to their

developmental level. They also added that students of this age did not want to criticize them because of their positive feelings for the classroom teacher. A teacher pointed out that:

You can't show me one student who doesn't like a primary school classroom teacher. This age group is too young, and it is not right to ask them to comment on effective teacher characteristics. Students may like their teacher not to teach anything, give less homework, or chat in the classroom. Or vice versa, he may not like his teacher that day because of a disciplinary event happening in the classroom at that moment. If you ask something about his teacher at that moment, he says negative things. Therefore, I do not think that the student can evaluate objectively. (T14 thirteen years of experience)

Bana ilkokul sınıf öğretmenini sevmeyen bir tane öğrenci gösteremezsin. Bu yaş grubu çok küçük ve onlardan etkili öğretmen özellikleri hakkında yorum istemek doğru değil. Öğrenciler, öğretmenini bir şey öğretmediği halde az ödev veriyor ya da sınıfta sohbet ediyor diye sevebilir. Ya da tam tersi o an sınıfta gerçekleşen bir disiplin olayı nedeniyle o gün öğretmenini sevmeyebilir. O anda öğretmeni hakkında bir şey sorarsanız, olumsuz şeyler söyler. Bu nedenle öğrencinin nesnel olarak değerlendirebileceğini düşünmüyorum. (T14 on üç yıllık tecrübe)

Another teacher noted that:

[...] The primary school student himself does not know what it means to evaluate. No student of this age says bad things about their teacher. They always make positive comments. Even though I have been giving homework for a week, students can say that "my teacher cares about me." (T16 six years of experience)

[...] İlkokul öğrencisi kendisi daha değerlendirmenin ne demek olduğunu bilmiyor. Bu yaştaki hiçbir öğrenci öğretmeni için kötü demez. Her zaman olumlu yorumlar yaparlar. Bir hafta boyunca ödev vermeme rağmen öğrenciler "öğretmenim benimle çok ilgileniyor ve beni önemsiyor" diyebilir. (T16 altı yıllık tecrübe)

Few teachers (T6, T13, T20) and principals (P3, P7) stated that parents could evaluate the teachers, but almost all the participants said that parents should not evaluate teachers. In most of these opinions, it was noted that it was not correct for parents to evaluate teachers according to what they heard from their children without knowing the teacher in the classroom. The participants also added that parents were insufficient to evaluate the teacher and cannot make an objective evaluation. One of the teachers explained the drawbacks of being evaluated by parents and students as follows:

Parents don't know this school like us. The person who evaluates me should know this school very well. The evaluator should be someone who knows what we are doing in the classroom. For example, my student's mother works in a bank. How will a banker evaluate me? The evaluator should observe me in the classroom. (T16 six years of experience)

Ebeveynler bu okulu bizim gibi bilmiyorlar. Beni değerlendirecek kişi bu okulu çok iyi tanımalıdır. Değerlendirici, sınıfta ne yaptığımızı bilen biri olmalıdır. Örneğin öğrencimin annesi bir bankada çalışıyor. Bir bankacı beni nasıl değerlendirecek? Değerlendirici beni sınıfta gözlemlemelidir. (T16 altı yıllık tecrübe)

4.1.3.1.6 Recommendations for Evaluation Method

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the evaluation methods are given in Table 4.11.

Table 4.11

Codes for The Evaluation Method

Theme 1. Evaluation Process
Evaluation Method
Observations
Debriefing after Observation
Self-evaluation

As previously explained, the participants stated that the problems related to the observations were the lack of systematic observations at school, the teachers not being informed about the results of the observations, and the lack of evaluation criteria. When the participants were asked how to get the evaluation data, most of the teachers ($f=18$), principals ($f=8$), and experts ($f=7$) stated observation forms should be used as data collection tools. The participants also added that using observation results as the source for the evaluation would be the most effective one to determine a teacher's effectiveness. They also stated that it was appropriate to provide the necessary data from the multiple observations. For example, one of the teachers said:

I think observation is crucial, and I will be happy to have observed my class many times. Reports of students with low success in exams are constantly sent to us, and it is

Gözlemin çok önemli olduğunu düşünüyorum ve sınıfımı birçok kez gözlemlenmesinden mutluluk duyarım. Sınavlarda başarısı düşük olan öğrencilerin raporları sürekli olarak

thought that we do nothing for these children. I would love them to come to my class and see how much effort I put into them. (T4 ^{three} years of experience)

bizlere gönderiliyor ve bu çocuklar için sanki hiçbir şey yapmadığımız düşünülüyor. Sınıfıma gelip onlara ne kadar emek harcadığımı görmelerini çok isterim. (T4 ^{üç} yıllık tecrübe)

Another teacher also pointed out that:

Sometimes I go to a friend who has a lesson in the same corridor and asks questions about a topic I can't explain. I ask how she describes the subject. It helps a lot. I wish I could enter her class as an observer. [...] I would like to make multiple observations with the help of a well-structured form. I must look at the criteria to give her feedback and understand what I missed while teaching the same topic. (T19 ^{three} years of experience)

Bazen aynı koridorda dersi olan bir arkadaşımın yanına gidiyorum ve ona anlatamadığım bir konu hakkında sorular soruyorum. Konuyu o nasıl anlatıyor diye soruyorum. Çok yardımcı oluyor. Keşke sınıfına bir gözlemci olarak girebilsem. [...] bu gözlemleri de iyi yapılandırılmış bir form yardımıyla yapmak isterim. Ona geri dönüt verebilmek ve aynı konuyu öğretirken neyi kaçırdığımı anlayabilmek için ölçütlere bakmalıyım. (T19 ^{üç} yıllık tecrübe)

While some teachers (T1, T2, T4, T6, T10, T11, T12, T15) and experts (E1, E2, E3, E6, E7) stated that observation was the only method that can be used to evaluate the instruction if the teacher trusts to the process of evaluation and evaluator. The principals (P1, P3, P4, P5, P9) explained that they should perform classroom observation whenever they could to understand teachers. One of the teachers had brought the following suggestions for the observation process to be effective:

Observation should be done, of course, but someone should explain why this observation was made. Suppose I find observation is made to see my weaknesses and use these weaknesses as evidence to make sanctions. In that case, I behave very differently in the classroom or prepare the children for lessons because there is a chance that something terrible will happen in the end. Do you know when I act like myself during the observation? When I trust this observation process. I then believe that observation is for my development

Gözlem elbette yapılmalıdır, ancak birisi bu gözlemin neden yapıldığını açıklamalı. Gözlem eksikliğimi bulmak ve bu eksikliği kanıt olarak kullanarak yaptırım uygulamaksa, o zaman sınıfta çok farklı davranırım veya çocukları derslere hazırlarım çünkü sonunda başıma kötü bir şey gelme ihtimali var. Gözlem sırasında ne zaman kendim gibi davranırım biliyor musun? Bu gözlem sürecine güvendiğimde. O zaman gözlemin gelişimim için olduğuna ve nesnel olarak

and is done objectively. (T13 twelve years of experience) yapıldığına inanırım. (T13 on iki yıllık tecrübe)

One of the experts explained as:

When conducting an observation, the teacher is concerned about the objectivity of the observation and the use of observation results. Teachers should be told why this observation was made and what to consider in the observation before the process begins. She should fully trust the process. It is necessary to give immediate feedback on the strengths and deficiencies identified after the observation. The teacher should not worry about how the observation went, for example (E2 six years of experience) Bir gözlem yapılırken, öğretmen gözlemin nesnelliği ve gözlem sonuçlarının kullanımı konusunda endişe duyar. Öğretmenlere bu gözlemin neden yapıldığı ve süreç başlamadan önce gözlemde nelere dikkat edileceği söylenmelidir. Bu sürece tamamen güvenmelidir. Gözlemden sonra belirlenen güçlü ve eksiklikler hakkında anında geribildirim vermek lazım. Öğretmenin gözlem nasıl geçti acaba diye düşünüp kaygılanmamalı mesela (E2 altı yıllık tecrübe)

Participants (T2, T3, T5, T9, T11, T12, T15, T19, P1, P3, P6, P9, E1, E4, E5, E6) stated making observation was the only way to understand the classroom practices totally and making multiple observations is the necessary process of conducting objective teacher evaluations. One of the teachers explained the importance of using different observation results:

Some days we cannot get the efficiency we want in the classroom. This efficiency may depend on me or the situation of the students. For example, it is impossible to teach in the classes in the week before the holiday, because children want to go on holiday as soon as possible and do not listen. It would be wrong for the observer to give me feedback by observing only that day. Therefore, observations should be made at different times throughout the year. (T17 fifteen years of experience). Bazı günler sınıfta istediğimiz verimi alamıyoruz. Bu bana veya öğrencilerin durumuna bağlı olabilir. Örneğin, tatilden önceki haftadaki sınıflarda ders yapmak mümkün değildir, çünkü çocuklar mümkün olan en kısa sürede tatile gitmek isterler ve dersi dinlemezler. Gözlemcinin sadece o günü gözlemleyerek bana geribildirim vermesi yanlış olur. Bu nedenle yıl boyunca farklı zamanlarda gözlem yapılmalıdır. (T17 on beş yıllık tecrübe).

One teacher explained the importance of using observations results as follows:

In my old school, my principal observed the teachers very often, and I would not be disturbed because I was very used to the process. The results of the observation aimed to improve us professionally. The observations were already in their natural flow. Sometimes my principal gave me the right to choose the course to be followed. Both written and verbal feedback were given. This feedback was given immediately so that we wouldn't wait. I never felt uneasy. I learned a lot to improve myself. Here, observation should provide us with this trust. If I trust the observer, I will be comfortable with the observation. (T12 seven years of experience)

Eski okulumda yöneticilerim öğretmenleri çok sık gözlemlerlerdi ve hiç de rahatsız olmazdım çünkü sürece çok alışkındım. Gözlemin sonuçları bizi profesyonel anlamda geliştirmeyi amaçlıyordu. Gözlemler doğal akışında oluyordu zaten. Müdürüm bana bazen gözlemlenecek dersi seçme hakkı da veriyordu. Hem yazılı hem de sözlü geri bildirim alıyordum. Bu dönüt hemen verilirdi yani beklemezdik. Hiç huzursuz hissetmedim. Kendimi geliştirmek için çok şey öğrendim. Burada da gözlem bize bu güveni vermelidir. Eğer gözlemciye güvenirsem, gözlem konusunda rahat olurum. (T12 yedi yıllık tecrübe)

Moreover, most of the teachers ($f=16$), principals ($f=7$), and experts ($f=5$) stated that the interviews after these observations, which were used for debriefing, would be essential. Teachers said that they could be evaluated more objectively by interviews made after the observation, and they could reduce the misunderstandings caused by the observer in the classroom. On the other hand, principals stated that asking teachers what happened in the classroom was an effective way to understand teachers' instructional behaviors objectively. Experts also indicated as debriefing after the observation was an indispensable part of interpreting the observation data, especially for classroom teachers, to fully understand the behaviors of these teachers who try to teach children at a very young age. Some of the participants (T1, T2, T4, T9, T12, T14, T13, T17, T20, P1, P2, P3, P4, P8, E1, E2, E3, E4, E6, E7) explained as especially the interviews before and after the observation would help make more accurate evaluations about the process and getting individual feedback. Some of them (T1, T3, T4, T5, T7, T18, P2, P3, P4, P5, E1, E2, E4) stated that these interviews would allow them to explain their ideas and express themselves. For example, one of the teachers said:

For example, that day, one of my students was ill and had trouble listening to me. Sometimes I say you can put your head on the table and just listen. The observer may think that my management in the classroom is weak because apparently, I am not interested in this student or ask questions to him. In this case, after the observation, I may have the opportunity to explain the status of this student only if asked. Otherwise, wrong assessment is made about classroom management. (T2 five years of experience)

Örneğin, o gün, bir öğrencim hastaydı ve beni dinlemekte zorlanıyordu. Bazen başını masanın üstüne koyabilir ve sadece dinleyebilirsin diyorum. Gözlemci sınıftaki yönetimimin zayıf olduğunu düşünebilir, çünkü görünüşte bu öğrenciyle sınıfta ilgilenmiyorum ya da ona sorular sormuyorum. Bu durumda, gözlemden sonra, sadece bana sorulursa bu öğrencinin durumunu açıklama fırsatına sahip olabilirim. Öteki türlü sınıf yönetimi hakkında yanlış değerlendirme yapılır. (T2 beş yıllık tecrübe)

One of the principals explained as:

I know a lot of teachers in this school and am very close friends with some of them. For example, when I observe behavior in this teacher's class that I think she could never do, I can ask why she did this after observation or ask questions to understand the behavior of a teacher I never knew. (P7 twelve years of experience)

Bu okulda çok fazla öğretmen tanıyorum ve bazılarıyla çok yakın arkadaşım. Örneğin, bu öğretmenin sınıfına girdiğim ve asla yapamayacağını düşündüğüm bir davranış gözlemlediğim zaman gözlemden sonra bunu neden yaptığını sorabilirim ya da hiç bilmediğim bir öğretmenin davranışını anlayabilmek için sorular sorabilirim. (P7 on iki yıllık tecrübe)

One of the experts pointed out that:

We make observations, take notes and leave the class during the evaluation process. It would be wrong to take the observation form and start reporting immediately. First, we must share what we observe with the teacher. These children are very young, and therefore I may have a false impression of the teacher. We should not blame the teacher for not being able to answer the question in the observer's mind. This process is best supported by asking and listening to the teacher after the observation. (E4 four years of experience)

Değerlendirme sürecinde gözlemler yapar, not alır ve sınıftan ayrılırız. Gözlem formunu alıp hemen raporlamaya başlasak yanlış olur. İlk olarak, gözlemlediğimiz şeyi öğretmenle paylaşmalıyız. Bu çocuklar çok küçük ve bu nedenle öğretmen hakkında yanlış bir izlenim sahibi olabilirim. Gözlemcinin zihnindeki soruyu cevaplayamadığı için öğretmeni suçlamamalıyız. Bu en iyi gözlem sonrası öğretmene sorarak ve dinleyerek desteklenir. (E4 dört yıllık deneyim)

Participants could not reach a consensus on using self-assessment in teacher evaluation. Most of the teachers ($f=16$) argued that everyone who properly fulfills the teaching profession, especially primary school teachers, had an opinion about themselves. They also noted that a teacher could best evaluate themselves by using self-evaluation tools. A teacher expressed her views on the use of self-assessment in the teacher evaluation process as follows:

Unfortunately, although self-assessment forms are essential measurement tools, they are not widely accepted. We couldn't teach this to our students either. In fact, self-esteem and self-control are crucial. Perhaps it is the tool that enables a person's development. We are afraid of these forms because we do not correctly apply and evaluate self-evaluation. But I think as a teacher, we must first evaluate ourselves and get used to these forms. (T18 ten years of experience)

Ne yazık ki, öz değerlendirme formları çok önemli ölçüm araçları olmasına rağmen, pek kabul görmezler. Bunu öğrencilerimize de öğretmedik. Aslında, benlik saygısı ve benlik kontrolü çok önemlidir. Belki de bir kişinin gelişimini sağlayan araçtır. Bu formlardan korkuyoruz çünkü öz değerlendirmeyi doğru bir şekilde uygulamıyor ve değerlendirmiyoruz. Fakat bence öğretmen olarak önce kendimizi değerlendirmeli ve bu formlara alışmalıyız. (T18 on yıllık tecrübe)

In addition to these opinions, most principals ($f=6$) mentioned that the self-assessment could be used if the self-assessment form was prepared accurately by setting clear criteria and supported by data collected through other tools. Experts ($f=4$) stated that they wouldn't favor self-evaluation if the necessary training was not taken, and the evaluation process was not managed correctly. Experts (E2, E6, E7) who do not find it appropriate to include self-evaluation results in the evaluation stated that teachers' self-perception was much higher than it should be and was not easy for the teacher to evaluate themselves objectively. One of the principals stated:

When I first thought of it, I was a little hesitant to say self-evaluation. However, I think it would be effective for the teacher to evaluate himself objectively. I think the teacher uses the form objectively if the criteria are prepared very clearly. But it should not be used alone. It should be considered in conjunction with what

İlk aklıma geldiğinde öz-değerlendirme demekten biraz tereddüt ettim. Ancak, öğretmenin kendini objektif olarak değerlendirmesinin etkili olacağını düşünüyorum. Ölçütler çok açık bir şekilde hazırlanmış olursa, öğretmen formu objektif olarak kullanır bence. Ama tek başına kullanılmamalıdır.

other observers have said. (P8 eight years of experience)

Diğer gözlemcilerin söyledikleri ile değerlendirilmelidir. (P8 sekiz yıllık tecrübe)

In contrast to these views, one expert stated:

I think this is an important data collection tool, but it is not suitable for classroom teachers to use it. Because they claim that they do everything exactly. We gave an evaluation form after the activities, and once again, I did not see that the teacher wrote something for herself. They say that the lack is due to the plan or student level. (E2 eight years of experience)

Bence bu önemli bir veri toplama aracı, ancak sınıf öğretmenlerinin kullanması doğru değil. Çünkü her şeyi tam olarak yaptıklarını iddia ediyorlar. Etkinliklerden sonra bir değerlendirme formu veriyoruz ve bir kez daha öğretmenin kendine yönelik bir şey yazdığını görmedim. Eksikliğin plandan ya da öğrenci seviyesinden kaynaklandığını söylüyorlar. (E2 sekiz yıllık tecrübe)

4.1.3.2 Recommendations for the Use of Evaluation Data

The recommendations reported by the participants are given under the headings of propriety, utility, and accuracy in this section. The participants also addressed how they would like to access the evaluation results.

4.1.3.2.1 Access to Evaluation Information

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts on access to assessment results are given in Table 4.12.

Table 4.12

Codes for The Access to Evaluation Results

<i>Theme 2. Use of Evaluation Data</i>
<i>Access to Evaluation Information</i>
Importance of confidentiality
Individual feedback (face to face)

One of the critical problems of this case is the lack of access to evaluation information. Almost all the teachers ($f=19$), most of the principals ($f=5$), and experts ($f=4$) stated that teacher information should have been kept confidential and their opinions about themselves should be explained only to the teacher herself. Besides, most of the teachers ($f=15$), principals ($f=8$), and all the experts ($f=7$) stated that it was essential to

inform the teachers individually through face-to-face interviews. A teacher explains the importance of individual meetings as follows:

[...] sometimes, our school principal gathers all the teachers and gives information about general disruptions, but nobody cares about the problem when this kind of information is provided. Nothing changes when the meeting ends. Of course, she does not say the shortcomings of a person in the middle of everyone but can give information privately, individually. (T12 seven years of experience)

[...] bazen okul müdürümüz tüm öğretmenleri toplar ve genel aksaklıklar ile ilgili bilgi verir, ancak bu tür bir bilgi verildiğinde hiç kimse yaşanan sıkıntıyı üstüne almıyor. Toplantı sona erdiğinde hiçbir şey değişmez zaten. Tabii ki herkesin içinde bir kişinin eksikliklerini söylemesin ama özel olarak bireysel olarak bilgi verebilir. (T12 yedi yıllık tecrübe)

One of the experts explained her view as follows:

I witnessed it once when I attended a meeting also including a topic concerning our department. The principal gathered all the 3rd-grade teachers and gave a public speech. Everyone was happy; nobody's face was sullen. No one takes over like that. When the meeting is over, the teacher continues to do what she knows. Feedback should be given individually and specifically to that teacher (E6 five years of experience)

Bir keresinde bizim birimi de ilgilendiren bir konu için toplantıya katıldığımda şahit olmuştum. Müdür tüm 3. sınıf öğretmenlerini toplamış genel konuşma yapıyor. Herkes memnun kimsenin yüzü asık değil. Böyle kimse üstüne alınmaz ki. Toplantı bittiğinde öğretmen bildiğini yapmaya devam ediyordur. Geri bildirim o öğretmenin özelinde ve bireysel verilmeli. (E6 beşi yıllık tecrübe)

4.1.3.2.2. Recommendations for Propriety

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts on providing propriety are given in Table 4.13.

Table 4.13

Codes for Providing Propriety

<i>Theme 2. Use of Evaluation Data</i>
<i>Propriety</i>
Balanced Evaluation
Professional interactions

Most of the teachers ($f=15$), principals ($f=6$), and all the experts ($f=6$) stated that the results of the assessments should have included both strengths and weaknesses. Participants explained their ideas by saying the importance of balanced evaluation, and that the evaluations made for this purpose would be more ethical. They also stated that this type of evaluation could allow the process to favor the evaluated people. One of the principals explained the importance of including both strengths and weaknesses in feedback sessions as follows:

One should know well what she is doing to contribute to her strengths and even guide other teachers in these strengths. In fact, when explaining the evaluation results, first the strengths and then the shortcomings can be explained. Weaknesses should not only be justified but also how these weaknesses can turn into strengths should be described. This process is the only way to trust evaluations.
(P6 sixteen years of experience)

Bir kişinin güçlü yönlerine katkıda bulunabilmesi ve hatta bu güçlü alanlardaki diğer öğretmenlere rehberlik edebilmesi için ne yaptığını iyi bilmelidir. Aslında, değerlendirme sonuçları açıklanırken önce güçlü yönler ve sonra eksik yönler açıklanabilir. Zayıf yönler sadece açıklanmamalı, aynı zamanda bu zayıf yönlerin nasıl güçlü yönlere dönüşebileceği de ele alınmalıdır. Değerlendirmelere güvenmenin tek yolu budur. (P6 on altı yıllık tecrübe)

Most principals ($f=7$) and experts ($f=7$) explained interacting professionally with the teachers to be evaluated could keep them from feeling anxious, and they believe that criticism made against them is made with respect. An expert (E2 six years of experience) stated as “*Treating in a bad or disrespectful manner to a teacher constantly to tell her mistakes, decreases her motivation and she starts to get scared (Hatalarını anlatmak için sürekli bir öğretmene kötü veya saygısız davranmak motivasyonunu azaltır ve öğretmen korkmaya başlar.)*” One of the principals explained as:

You can make any kind of criticism of the teachers, but how you do this is very important. Sometimes I must explain the mistakes she made to a teacher 30 years older than me. In this case, I do not set up command sentences. First, I need to explain why I have to explain this situation, what we can do to fix it, or why it's essential. This professional

Öğretmenlere her türlü eleştiriyi yapabilirsiniz, ancak bunu nasıl yaptığınız çok önemli. Bazen benden 30 yaş büyük bir öğretmene yaptığı hataları açıklamak zorunda kalıyorum. Emir cümleleri kurmuyorum. Önce bu durumu neden açıklamak zorunda olduğumu, bunu düzeltmek için neler yapabileceğimizi açıklamam gerekiyor veya bunu

communication should always be continued (P5 six-years of experience) yapmasının neden önemli olduğunu. Bu profesyonel iletişim her zaman devam ettirilmeli. (P5 altı yıllık tecrübe)

4.1.3.2.3 Recommendations for Utility

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts on providing utility are given in Table 4.14.

Table 5

Codes for Providing Utility

Theme 2. Use of Evaluation Data
Utility
Explicit criteria for the usefulness of the results or defensibility
Functional reporting

Most of the teachers ($f=15$), principals ($f=7$), and all the experts ($f=7$) stated that all the tools used should have been based on explicit criteria so that interpretation and judgment based on these criteria could make sense and, thus, an open and defensible assessment environment was created. Most of the teachers ($f=13$) and experts ($f=6$) also mentioned that the conclusions about the teacher's performance should be justified, and this could only be achieved by defining the criteria before the evaluation. They also stated that this was a way to prove to be accurate in evaluation. One of the experts (E4 four years of experience) explained as "If you do not have detailed and well-structured criteria, there is no way to explain or justify what you are evaluating (Detaylı hazırlanmış ve iyi yapılandırılmış kriterleriniz yoksa, neyi değerlendirdiğinizi açıklamanın veya değerlendirmeyi haklı çıkarmanın bir yolu yoktur.)" A teacher expressed the importance of evaluation by criteria and her suggestions about this subject as follows.

Evaluations should be criteria-based, and the results should inspire me. So, I should be able to develop it for myself with the help of these results. The criteria ultimately give clear information on what I should improve on performance. If general Değerlendirmeler ölçüt temelli olmalı ve sonuçlar bana ilham vermelidir. Yani bu sonuçlar yardımıyla kendime geliştirebilmeliyim. Ölçütlerde sonuçta performans konusunda neyi iyileştirmem gerektiğine dair net bilgiler verir. Hiçbir ölçütün

expressions are used when there is no criterion, I cannot object to what is said with this general expression because it is not clear how much of this general situation I have done and how much I could not do. For example, when students speak a lot, when I am told that classroom management is weak, I would like to ask them about what process of management in which lesson this activity was observed. I do some activities just to let the children speak. (T11 twenty years of experience)

olmadığında genel ifadeler kullanılırsa bu genel ifadeyle söylenenlere itiraz edemem çünkü bahsedilen bu genel durumun ne kadarını yaptım ne kadarını yapamadım belli değil. Örneğin, öğrenciler çok konuştuğunda, sınıf yönetiminin zayıf olduğu söylendiğinde, onlara tam olarak hangi yönetim süreci eksikti, hangi derste gözlemlendiğini sormak istiyorum. Sadece çocukların konuşmasına izin vermek için bazı etkinlikler yapıyorum. (T11 yirmi yıllık tecrübe)

One of the teachers stated:

We cannot change without measuring and evaluating. I believe there should be an assessment for teachers, but how it is done is very important. If this assessment cannot be made according to specific criteria, the decisions taken cannot be justified and have no meaning. (T18 eighteen years of experience)

Ölçmeden ve değerlendirmeden değişemeyiz. Değerlendirmenin olması gerektiğine inanıyorum, ancak nasıl yapıldığı çok önemli. Bu değerlendirme belirli kriterlere göre yapılamazsa, alınan kararlar gerekçelendirilemez ve hiçbir anlamı yoktur. (T18 on yedi yıllık tecrübe)

Participants determined functional reporting as a report including feedback and teacher growth plan based on multiple measurements. The majority of the teachers ($f=16$), principals ($f=7$), and experts ($f=6$) explained that reports should be given at the end of each education period to give teachers enough time for their development. They also stated that if the report is given at the end of the year, sufficient time would not be given for the teacher to complete her professional development by removing the deficiencies identified. They explained as it was essential because the results of the evaluations would be meaningless, and the evaluation would not have a practical value if the results were not presented. Some of the teachers (T5, T10, T15, T20) stated that the school could report at the end of the year. They explained that it would be challenging to report more frequently due to the high number of teachers and to collect enough data in a period. One of the experts explained her ideas about functional reporting as:

In particular, feedback should be given in a written report. As an expert, we explain to teachers that if you are not going to give feedback to the students, do not give your students homework because when you do not provide feedback, the student does not do homework again. The same is true for principals and teachers here. When a teacher does right or wrong, not being followed and not giving feedback gives the teacher the idea, "No one cares what I do anyway. Let me continue to do what I do." Such behaviors do not provide professional development. (E2 *six years of experience*)

Özellikle yazılı bir raporla geri bildirim verilmeli. Uzman olarak, öğretmenlere dönüt vermeyecekseniz öğrencilerinize ödev vermeyin diyoruz çünkü geri bildirimde bulunmadığınız zaman öğrenci tekrar ödev yapmıyor. Aynı durum burada yöneticiler ve öğretmenler için de söz konusudur. Bir öğretmenin doğru ya da yanlış yaptığında, takip edilmemesi ve geri bildirimde bulunulmaması, öğretmene "Nasıl olsa ne yaptığımı kimse umursamıyor ben de. Yaptığım şeyi yapmaya devam edeyim" fikrini verir. Bu tür davranışlar profesyonel gelişimi sağlamaz. (E2 *altı yıllık tecrübe*)

One of the principals explained as:

For example, as a result of evaluations in the first two months, we observed a significant deficiency in classroom management or teaching of the teacher. We have such a result; what do we do? Should we wait and give feedback at the end of the academic year? It would be unfair to wait until the end of the year for students in this class. The teacher should get feedback as soon as possible about this. Feedback should be given within a report at the end of the first semester at the latest. (P1 *thirteen years of experience*)

Örneğin, ilk iki aydaki değerlendirmeler sonucunda, öğretmenin sınıf yönetimi veya öğretmenliğinde önemli bir eksiklik gözlemledik. Elimizde böyle bir sonuç var peki ne yapacağız? Beklemeli ve eğitim-öğretim yılı sonunda mı geribildirim vermeliyiz? Bu sınıftaki öğrenciler için yıl sonuna kadar beklemek haksızlık olur. Öğretmen bu konuda en kısa zamanda geri bildirim almalıdır. Geri bildirim bir raporla en geç ilk dönemin sonunda verilmelidir. (P1 *on üç yıllık tecrübe*)

4.1.3.2.4 Recommendations for Accuracy

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts on providing accuracy are given in Table 4.15.

Table 4.15*Codes for Providing Accuracy*

<i>Theme 2. Use of Evaluation Data</i>
<i>Accuracy</i>
Bias Identification and Management
Equal evaluation regardless of teaching experience
Use of multiple data sources and method
Open communication
Positive environment

Almost all the teachers ($f=18$), principals ($f=7$), and all experts ($f=7$) mentioned that evaluation information should be obtained objectively, so it should not be based on bias identification meaning that decisions should not be made based on personal relations. They stated that evaluation results should ensure interpretations from the data collected about a teacher's performance and should not be open to misinterpretation. One of the teachers explained as:

There are teachers that I am very close to in school and those who do not like me very much. They may tend to interpret the behaviors they observe as good or bad as they are. In this case, the observed situation does not reflect what actually happened, which leads to incorrect evaluations. A classroom teacher who doesn't know me well can observe me. (T2 five years of experience)

Okulda çok yakın olduğum öğretmenler var ve beni çok sevmeyenler var. Gözlemledikleri davranışları gerçekte olduğumdan iyi ya da kötü olarak yorumlama eğiliminde olabilirler. Bu durumda, gözlemlenen durum gerçekte ne olduğunu yansıtmaz, bu da yanlış değerlendirmelere yol açar. Beni çok iyi tanımayan bir sınıf öğretmeni beni gözlemleyebilir. (T2 beş yıllık tecrübe)

One of the experts mentioned:

Ensuring validity is one of the most critical aspects of assessment and is more critical for teacher evaluation. Teachers should not be treated differently to ensure validity. They can be good friends at school or out of school, but they have to put this aside during the evaluation process. The best way to achieve this is not allowing

Geçerlik sağlanması, değerlendirmenin en önemli unsurlarından biridir ve öğretmen değerlendirmesi için çok daha önemlidir. Geçerliği sağlamak için öğretmenlere farklı davranılmamalıdır. Okulda veya okul dışında çok iyi arkadaş olabilirler, ancak değerlendirme sürecinde bunu bir kenara bırakmak zorundadırlar. Bunu başarmanın en iyi yolu, birbirine çok yakın olan

teachers who are very close to each other to observe each other. In this school, you know very well which teacher is very close friends. The observer should be able to observe the teachers without any bias. (E7 three years of experience)

öğretmenleri birbirini gözlemlemesine izin vermemek. Bu okulda hangi öğretmen çok yakın arkadaş çok iyi bilinir. Gözlemci, herhangi bir önyargı olmaksızın öğretmenleri gözlemleyebilmelidir. (E7 üç yıllık tecrübe)

Many teachers ($f=14$) complained that the experienced and less experienced teachers were treated differently in this school, and this situation was highly likely to be reflected in the evaluation results. The teachers stated that the behaviors of experienced teachers were tolerated based on the good relations of these people with the principals. On the other hand, the behavior of a newly started teacher was constantly examined and tried to be found incomplete. Participants suggested that regardless of their professional experience, it was important to evaluate all teachers by following the same process. One teacher explained this situation as:

Even my clothes were criticized when I first started this school, but another teacher was not charged for her outfit because she had a lot of experience. There is an evaluation in favor of experienced teachers in every subject. Therefore, assessment should be the same for all, irrespective of the teaching experience, and even the criteria should be the same. In this case, I think the observer has a lot of work to do and should treat teachers with different years of experience equally. (T15 four years of experience)

Bu okula ilk başladığımda kıyafetlerim bile eleştirildi, ancak başka bir öğretmen kıyafeti konusunda eleştiri almadı çünkü bu okulda çok fazla deneyime sahipti. Her konuda tecrübeli öğretmenlerin lehine bir değerlendirme söz konusu. Bu nedenle değerlendirme, öğretim deneyiminden bağımsız olarak herkes için aynı olmalıdır, hatta kriterler aynı olmalıdır. Bu durumda, gözlemcinin yapacak çok işi olduğunu ve farklı deneyime sahip öğretmenlere eşit davranması gerektiğini düşünüyorum. (T15 dört yıllık tecrübe)

Most of the teachers ($f=15$), all principals ($f=9$), and all the experts ($f=7$) stated that using multiple data sources was one of the important things to provide consistent indicators of teacher behaviors in a different classroom or out of classroom settings. They also added this would also be essential to provide reliable information and to be accurate in evaluation. For example, one of the teachers explained this:

This year the principal made an observation. I was really sick that

Bu yıl okul müdürü bir gözlem yaptı. O gün gerçekten hastaydım. Sınıfta

day. I couldn't perform the performance I wanted in the classroom. I explained this to the principal before class, and she understood me, but I think I am a good teacher, and I should have had another opportunity to show it. My teaching should not be evaluated by looking at a single performance. In fact, only observation should not be made. Okay, maybe I told my principal that I was sick, but I could not do this. (T3 fourteen years of experience)

istediğim performansını tam olarak yerine getiremedim. Bunu dersten önce okul müdürüne açıkladım ve o da beni anladı, ama bence ben iyi bir öğretmenim ve bunu göstermek için başka bir fırsatım daha olmalıydı. Öğretimim tek bir performansa bakarak değerlendirilmemelidir. Aslında, sadece gözlem de yapılmamalıdır. Tamam belki ben hasta olduğumu müdürüme söyledim ama söyleyemeye de bilirdim. (T3 on dört yıllık tecrübe)

One of the principals explained as:

Sometimes parents complain about the teacher, but I do not know what that teacher did in the lesson to give the right or defend the teacher. Currently, we are not allowed to observe teachers by attending classes. I just do not find it right to look at the classroom in front of the door. I should observe three or four times by following the class of teachers in the corridor where I work as assistant principal. So, my decision can be correct and meaningful. (P7 twelve years of experience)

Bazen velilerden şikâyet geliyor öğretmenle ilgili ama ben o öğretmenin derste neler yaptığını bilmiyorum ki veliye hak vereyim ya da öğretmeni savunayım. Şu anda, derslere katılarak öğretmenleri gözlemlememize izin verilmiyor. Sadece kapının önündeki sınıfa bakmayı doğru bulmuyorum. Müdür yardımcısı olarak görev yaptığım koridordaki öğretmenlerin sınıfına katılarak üç-dört kez gözlemleyebilmeliyim. Böylece verdiğim karar doğru ve anlamlı olabilir. (P7 on iki yıllık deneyim)

Most of the teachers (f=15), principals (f=6), and experts (f=6) stated findings should be presented in an environment that should allow the teacher to express herself without fear, defend herself when necessary, or accept the decisions and feedback with confidence. The participants stated that such a secure and open environment could be provided after the observations and in the reporting process. For example, one of the teachers explained:

I am disconcerted about this because a decision has been taken, and I am not given any right to speak. I can neither talk to the principals nor my head of the department. Evaluation

Ben bu konuda çok dertliyim çünkü bir karar alınıyor hiç söz hakkı verilmiyor bana. Ne yöneticilerle konuşabiliyorum ne de zümre başkanıyla. Değerlendirme

results should be explained face to face, and the teacher should defend herself when necessary. The teacher should be allowed to protect her right and express herself. She should not feel fear of expressing herself. (T4 ^{three} years of experience)

sonuçları yüz yüze açıklanmalı ve öğretmen gerektiğinde kendini savunmalıdır. Öğretmenin hakkını savunmasına ve kendini ifade etmesine izin verilmelidir. Öğretmen kendini ifade etmekten korkmamalıdır. (T4 ^{üç} yıllık tecrübe)

The school principal stated that she was aware of the importance of open communication with the following words

My perceptions based on what I saw in my observations may not be correct. So, I want to meet with each teacher to understand the whole of their behavior after the observation, but we are a crowded school, and this is not possible. This way, if the teacher has something to say about observation, I would like to listen to her to make a more meaningful decision. (P4 ^{seventeen} years of experience)

Yaptığım gözlemlerde gördüklerimden yola çıkarak yaptığım algılamalar doğru olmayabilir. Bu yüzden gözlemden sonra davranışlarının bütünü anlamak için her bir öğretmenle görüşmek istiyorum ama mümkün olmuyor çünkü çok kalabalık bir okuluz. Bu şekilde, öğretmenin gözlem hakkında söyleyecek bir şeyi varsa, daha anlamlı bir karar vermek için onu dinlemek isterim. (P4 ^{on} yedi yıllık tecrübe)

The majority of the teachers (f=12) and principals (f=6) stated that using positive and constructive language while giving feedback to teachers would enable them to approach more moderately what was explained about them. Only in such an environment would they find the courage to defend themselves. For example, a teacher used the following statements while explaining the importance of a positive communication environment.

When judgmental and punitive language is used, I feel nervous and always need to attack. In this school, you are expected not to respond to the criticisms made against you, accept, and continue. So, I always swallow what is said. But this is not me. However, I can express myself more comfortably in a different, more positive environment. (T15 ^{four} years of experience)

Yargılayıcı cezalandırıcı bir dil kullanıldığında kendimi gergin ve hep saldırmam gerekiyormuş gibi hissediyorum. Bu okulda da genelde size yapılan eleştirilere karşılık vermemeniz kabul edip devam etmeniz bekleniyor. Yani hep yutuyorum söylenenleri. Ama bu ben değilim. Oysaki farklı yani daha pozitif bir ortamda kendimi daha rahat ifade ederim. (T15 ^{dört} yıllık tecrübe)

A vice-principal explained this situation as follows:

When I was a teacher, I could not express myself when talking to my principal. Now, for example, I observe that teachers come to me shyly. I always try to be positive to comfort them, and now they can open up to me more easily. While the teacher hears the comments about herself, the person on the other side should be positive so that teacher feels comfortable and express herself. (P2 twenty-two years of experience)

Ben öğretmenken de bu böyleydi yöneticimle konuşurken çekinir kendimi ifade edemezdim. Şimdi mesela öğretmenlerin yanıma çekinerek geldiğini gözlemliyorum. Onları rahatlatmak için hep pozitif olmaya çalışıyor ve artık bana daha rahat açılıyorlar. Öğretmen kendiyle ilgili yorumları duyarken de karşı taraftaki kişi olumlu olmalı ki öğretmen bir rahat etsin kendini açıklasın. (P2 yirmi iki yıllık tecrübe)

4.1.4 Participant Opinions for Professional Development Processes (Research Question 4)

The data to answer the fourth research question regarding the participants' opinions in terms of professional development was gathered through interviews conducted with teachers, principals, and experts working in that school. For this purpose, the participants were asked about the current state, strengths, and weaknesses of the professional development activities carried out in this school. In addition, the participants were also asked for recommendations for improvement of these professional development activities.

4.1.4.1 Planning the Process of Professional Development

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts on the planning process of the trainings for the professional development of teachers in the school are given in Table 4.16.

Table 4.16

Codes for The Planning Process of Professional Development

<i>Theme 1. Planning Training</i>
The needs of the majority of teachers
Educational issues on the agenda
Opinions of school experts

Almost all the teachers ($f=19$), all the principals ($f=9$), and experts ($f=7$) stated current professional development training was being planned based on the needs of most teachers. Participants stated that as opinions of teachers and experts were taken before each in-service training period, and training was planned for the most preferred subjects. For example, one of the teachers explained:

They give us a form before each in-service training period begins, and we write the five titles that we want to receive the most training. Then I don't know who evaluates them, but the most-posted title is selected. Then, finally, a seminar on that topic is given (T18 ten years of experience).

Bizlere her hizmetçi eğitim dönemi başlamadan önce bir form verirler ve bu forma en çok eğitim almak istediğimiz beş başlığı yazarız. Sonra bunları kim değerlendiriyor bilmiyorum ama en çok yazılan başlık seçiliyor. O başlıkla ilgili seminer veriliyor. (T18 on yıllık tecrübe)

One of the principals explained:

In some in-service training sessions, we ask teachers and experts to write what training they want. They write their requests. The people who can train in the fields which are written mostly by the teachers or chosen are determined. In our school, this subject is shared with teachers in the form of general training. (P4 seventeen years of experience)

Bazı hizmet içi eğitim dönemlerinde öğretmenlere ve uzmanlara sorarız ne ile ilgili eğitim almak istersiniz diye. Onlar da isteklerini yazarlar. Öğretmenler tarafından en çok ne yazıldıysa yani ne seçildiyse o alanda eğitim verebilecek kişiler belirlenir. Okulumuzda genel eğitimler şeklinde bu konu öğretmenlerle paylaşılır. (P4 on yedi yıllık tecrübe)

Some teachers ($f=7$), principals ($f=5$), and experts ($f=4$) stated that the principals also determined the training to be given during the in-service training period about the issues that became popular or raised in the current academic year. For example, one of the teachers explained:

There are things like that, for example, a person is trendy at that time, writes about education, and can come to our school and give a seminar. For example, the founder of the n-brain once came. Michio Kaku came in once, and it was quite a big event. Or we are also taking a seminar on a popular topic that we

Aslında şöyle şeyler olduğu da oluyor mesela bir kişi o dönemde çok popüler, eğitim hakkında yazılar yazıyor, o kişi de okulumuza gelip seminer verebiliyor. Mesela n-beyinin kurucusu gelmişti bir keresinde. Michia Kaku geldi bir kere de ve baya büyük bir organizasyondur aslında. Ya da televizyonda ve eğitimle ilgili

read or watch on television and education. (T14 thirteen years of experience). *okuduğumuz ya da izlediğim popüler bir konu hakkında da seminer alıyoruz. (T14 on üç yıllık tecrübe)*

Some principals (f=4) and experts (f=5) also stated that the opinions of the school experts were received about in-service training. For example, one of the experts noted this practice in professional development seminars as follows:

Principals, especially our general principal, care about our ideas based on our experience in education and our research in this area. We also work with teachers, and I think we know them well. Not always, in some periods, a subject that we have determined is shared with the teachers by a person we have selected. (E3 ten years of experience) *Yöneticiler özellikle de genel müdürümüz bizlerin eğitim alınındaki tecrübelerine ve bizim bu alanda yaptığımız araştırmalara dayanarak fikirlerimizi önemserler. Bir de biz öğretmenlerle çalışıyoruz ve onları iyi tanıyoruz bence. Her zaman olmasa da özellikle bazı dönemlerde bizim belirlediğimiz bir konu yine bizim belirlediğimiz bir kişi tarafından öğretmenlerle paylaşılır. (E3 on yıllık tecrübe)*

One of the principals detailed this practice in professional development seminars as follows.

The one thing our school is lucky to have experts here. There are both curriculum development and measurement and evaluation experts in our school. These experts are constantly working with teachers. I know they are sometimes asked to determine the training topics. I think they choose good subjects for teachers. (P8 seven years of experience) *Okulumuz şanslı olduğu bir konu burada uzmanların olması. Okulumuzda hem program geliştirme hem de ölçme değerlendirme uzmanları var. Bu uzmanlar öğretmenler ile sürekli çalışıyor. Hangi konuda eğitim verilsin diye bazen onlara sorulduğunu biliyorum. Onlarda bence güzel konular seçiyorlar öğretmenler için. (P8 yedi yıllık tecrübe)*

4.1.4.2 Perceived Strength and Weaknesses of Professional Development Practices

The participants were asked about the strength and weak aspects of the areas built in this school to ensure the professional development of teachers. The participants expressed their opinions about the elements that need to be developed rather than the strengths. The codes that emerged as a result of the analysis of the opinions of

teachers, principals, and experts on the strengths and problems of the professional development process are given in Table 4.17.

Table 4.17

Codes for The Strengths and Problems of Professional Development Processes

<i>Theme 2. Strengths of the professional development process</i>
Supportive environment
<i>Theme 3. Problems with the professional development process</i>
Training in large groups
Failure to evaluate the impact of training
Repetition of training of the same content

Some teachers ($f=11$), principals ($f=5$), and experts ($f=4$) stated that there was a supportive environment for teachers to receive training in their schools as strengths. However, the participants, especially teachers, had noted that this situation has decreased in recent years. Some principals (P1, P3, P4, P7) pointed out that the school supported teachers' education quite a lot, but the demand for training was not very high. Experts (E1, E2, E3, E6) stated that the school supported training for the teachers but that the teachers participated as listeners, the number of teachers who wanted to make presentations was quite a few, and that the teachers only participated in the training when they were obliged to make presentations. For example, a teacher expressed her opinion as follows:

For example, I want to attend many pieces of training, and the school board says, let me know the training you want to join. In fact, the school supports us, but we are class teachers, and we cannot leave the class and go to education [...] In the past, the school would pay all the education costs, but now, for example, it would cover the participation fee, but the travel costs belong to us. (T8 seven years of experience)

Şimdi mesela ben pek çok eğitime katılmak istiyorum ve okulda aslında bana katılmak istediğiniz eğitimleri söyleyin diyor. Aslında okul bizi destekliyor ama biz sınıf öğretmeniyiz ve sınıfı bırakıp eğitime gidemiyoruz [...] Eskiden okul tüm eğitim masrafını karşılardı şimdi mesela katılım ücretini karşılıyor ama yol masrafı bize ait oluyor. (T8 yedi yıllık tecrübe)

A principal explained this situation as follows:

Every year we say that tell us the training you want. They just need to tell us the reasons. However, the demand is so low that teachers do not approach to receive education during the school period, and they benefit from general education coming to school during the in-service education period. (P8 seven years of experience)

Her yıl diyoruz ki istediğiniz eğitimleri bize söyleyin. Bu okulda tüm öğretmenlerin istediği eğitime katılma şansı var. Bize gerekçeleriyle bildirmeleri gerekiyor sadece. Ancak talep o kadar az ki. Öğretmenler dönem içinde eğitim almaya pek yanaşmıyor ve hizmetçi eğitim döneminde ise okula gelen genel eğitimlerden yararlanıyorlar. (P8 yedi yıllık tecrübe)

The views of an expert on this subject are as follows:

I think teachers are supported in terms of education in this school. But if the teacher is going to go to this training, she is only a listener. There are very few teachers who want to explore and present something. If she's going to give a presentation, this happens because the school wants to. In fact, I think she attends a lot of training because the attendance is being checked. (E6 five years of experience)

Öğretmenler bu okulda eğitim anlamında destekleniyor bence. Ama öğretmen bu eğitime gidecekse sadece dinleyici oluyor. Bir şey araştırıp sunmak isteyen çok az. Bir de sunum yapacaksa mesela okul yap dedi diye yapıyor. Hatta bence pek çok eğitime yoklama alındığı için katılıyor. (E6 beş yıllık tecrübe)

More than half of the teachers ($f=11$), most of the principals ($f=5$), and experts ($f=6$) stated that performing educational activities to ensure professional development in this school in large groups reduced the effectiveness of the training. The participants noted that the reasons for this situation were the ignorance of individual needs, the lack of stability of the training, and the fact that the participants were only listeners and teachers could not participate actively. One teacher stated the weakness of this situation:

Principals and education experts tell us to make students active in the classroom and learn by doing. Why do we sit in a big room that is very crowded and listen for hours in training? I honestly remember very little of what I learned when I left training. It really doesn't help.

Hani yöneticiler ve eğitim uzmanları bize diyor ya öğrencileri aktif kılın sınıfta, onların yaparak öğrenmesine izin verin. Peki, biz niye eğitimlerde koca bir salonda sıkış tepiş oturup saatlerce dinleme yapıyoruz. Ben açıkçası eğitimden çıkınca öğrendiklerimden çok azını

Why can't we get training in small groups for whatever we need? (T14 thirteen years of experience)

hatırlıyorum. Gerçekten bir faydası olmuyor. Neden küçük gruplar halinde ihtiyacımız neyse ona göre eğitim alamıyoruz? (T14 on üç yıllık tecrübe)

Another teacher explained the weakness of training in large groups:

The most effective training I attended was drama and mathematics literacy. Do you know why? Because I was in the lead. Groups were 10-15 people. You can't do this with a crowded layout. Yes, maybe it can be inspiring in some seminars, but I don't know how to apply what I learned in my classroom. (T16 six years of experience)

Benim katıldığım en etkili eğitimler drama ve matematik okuryazarlığı eğitimleriydi. Niye biliyor musun? Çünkü başrolde ben vardım. Gruplar 10-15 kişiydi. Bunu kalabalık bir düzenle yapmanız mümkün değil. Evet belki seminer şeklinde yapılan bazı eğitimlerde ilham verici olabiliyor ama bu eğitimde öğrendiğimi kendi sınıfımda nasıl uygulayacağımı hiç bilmiyorum. (T16 altı yıllık tecrübe)

Regarding the weaknesses of existing professional development practices, most of the teachers ($f=16$), principals ($f=6$), and all the experts ($f=7$) indicated that the effectiveness of the professional development practices was not evaluated effectively. Although the effectiveness of training was questioned with a short questionnaire, participants explained that there was no clear assessment of what teachers changed or did differently due to the training or whether the subjects learned in training were applied effectively. For example, one of the teachers explained this weakness as follows:

We receive training and are expected to use what we have learned from this training in the classroom, that is, to reflect these practices into the teaching environment. Yes, this expectation is very appropriate, but the principal has never asked me what this training has added to you. I think we listen to education and leave what we have learned there. Then we speak very rarely, even in the department, about how to apply it. (T15 four years of experience)

Eğitimler alıyoruz ve sınıfta bu eğitimlerden öğrendiklerimizi kullanmamız yani bu uygulamaları öğretim ortamına yansıtmamız bekleniyor. Evet bu beklenti çok yerinde ama yönetim bana bu eğitim sana ne kattı diye hiç sormadı ki. Bence eğitimi dinliyoruz ve öğrendiklerimizi orda bırakıyoruz. Sonrasında zümrede bile çok nadir konuşuyoruz bunu nasıl uygulayalım diye. (T15 dört yıllık tecrübe)

One of the principals explained this weakness as follows:

As a matter of fact, sometimes we ask the teachers to see if the education has worked and sometimes even, they say that it was a very helpful training. But we don't do it officially. I thought we should do it now. (P7 twelve years of experience)

Aslına bakacak olursanız evet öğretmenlere bazen soruyoruz eğitim işe yaradı mı diye hatta bazen kendileri gelip bize eğitim çok faydalı oldu diyor. Ama bunu resmi bir şekilde yapmıyoruz. Şimdi düşündüm de yapmamız gerekir aslında. (P7 oniki yıllık tecrübe)

One expert stated the weakness as mentioned below:

Drama education, for example. All teachers received drama training for days and in small groups. No one asked these teachers, "was your drama training enough? What else is needed?". Or she doesn't go into class and observe a drama lesson to understand that this training works. (E6 five years of experience)

Drama eğitimi örneğin. Bütün öğretmenler drama eğitimi aldı günlerce ve küçük gruplar halinde. Kimse gidip de bu öğretmenlere aldığınız drama eğitimi yeter miydi, başka nelere ihtiyaç var demiyo ya da sınıfa girip bir drama dersi gözlemlemiyor ki bu eğitimin işe yaradığını anlasınlar. (E6 beş yıllık tecrübe)

Some teachers (f=7) and most of the experts (f=5) believed that a piece of training was repeated. On the other hand, there was no opinion among the principals regarding the repetition of the training. Teachers and experts stated that this was due to the lack of assessment after the training and the inability to identify the needs of teachers accurately. One of the experts (E3 ten years of experience) explained, "You cannot give the right training unless you determine the need correctly. So, some ineffective ones are being repeated (İhtiyacı doğru belirlemeden doğru eğitimi veremezsiniz. Bu yüzden etkisiz olanlar tekrarlanıyor)" One of the teachers (T14 thirteen years of experience) interpreted as " It is not known whether the training was useful or what we still lack so that training cannot be organized effectively and repeated unfortunately (Eğitimin faydalı olup olmadığı ya da halen eksiklerimizin olup olmadığı bilinmediği için eğitimler etkin bir şekilde organize edilemiyor ve maalesef gereksice tekrarlanıyor.)"

What an expert express on this subject is remarkable.

We regularly work with teachers every week. From the dialogues in the working groups, the plans prepared by the teachers, and the questions they ask

Biz her hafta öğretmenler ile düzenli çalışıyoruz. Çalışma gruplarında yaşanan diyaloglardan, öğretmenlerin hazırladığı planlardan ve bize sordukları sorulardan

us, we realize what they need. None of these training was aimed at meeting these needs. Moreover, the effectiveness of this training is not evaluated, and inadequate training is sometimes repeated. Now we can say that in this case, there was a severe mistake in planning the training here. (E5 two years of experience)

onların neye ihtiyacı olduğunu fark ediyoruz. Yapılan eğitimlerden hiçbiri bu ihtiyaçları gidermeye yönelik değildi. Kaldı ki zaten bu eğitimlerin etkililiği de değerlendirilmiyor ve kötü bir eğitim bazen tekrar ediliyor. Şimdi bu durumda burada eğitimler planlanırken ciddi bir yanlış olduğunu söyleyebiliriz." (E5 iki yıllık tecrübe)

4.1.4.3. Needs for Professional Development Practices

The codes that emerged as a result of the analysis of the opinions of teachers, principals, and experts regarding the needs for professional development practices are given in Table 4.18.

Table 4.18

Codes for The Professional Development Needs

<i>Theme 4. Needs of the professional development process</i>
<i>Data Source</i>
Determining the needs based on teacher evaluation
Measuring the effectiveness of training
Access to different instructional processes

The participants were asked to make suggestions about the weaknesses they mentioned before. Mainly participants suggested that while planning professional development activities, teacher evaluation results should be used. In addition, there should be a practical evaluation after the training through classroom observations and interviews. Finally, participants also stated that knowing the practices outside their classes would provide professional development.

Most of the teachers ($f=15$), principals ($f=8$), and all the experts ($f=7$) stated that the needs of teachers could be determined with an effective and reliable teacher evaluation system to ensure professional development. When the needs were determined in this way, the participants stated that the planning of the training to be included in the professional development would be effective in meeting the individual needs of the teaching, and the teacher would be allowed to develop themselves in the fields that they

really believe, and that the vocational development training would improve the quality of the teaching. A teacher explained her suggestion as follows:

Sometimes I cannot even determine in what sense I am strong or weak in my classroom practices. Now, even if I don't know it, how can someone who doesn't know me tell me the training I need. These weaknesses only occur when one comes and observes me. That is why we have been talking from the beginning, or in fact, as I said in this meeting, if there was such a system, there would be many people determining my weaknesses, and in the end, how to improve myself can be planned. (T7 fifteen years of experience)

Bazen sınıfta yaptığım uygulamalarda kendimin hangi anlamda güçlü ya da zayıf olduğunu ben bile belirleyemiyorum. Şimdi bunu ben bile bilmezken beni hiç tanımayan biri nasıl bilebilir ve bana ihtiyacım olan eğitimi söyleyebilir. Bu ancak bir beni gelip gözlemlediğinde ortaya çıkar. O yüzden hani başından beri konuşuyoruz ya aslında bu görüşmede söylediğim gibi bir sistem olsa benim eksikliği belirleyen bir sürü kişi olur ve sonunda kendimi nasıl geliştirebileceğim planlanabilir. (T7 on beş yıllık tecrübe)

One of the experts explained her suggestion as:

This is actually the case in school exams. If you develop the correct measurement tool, you will measure what you want to measure correctly. Determining the teacher's needs is, in fact, can be done through an accurate teacher evaluation system. The teacher also accepts the weaknesses as a result and will be willing to receive training. The teacher doesn't take anything from outside. Even if you offer the most successful education, they can be bored or unwilling. Development happens with the proper assessment. How can we improve the weakness if we cannot detect it? (E2 six years of experience)

Bu aslında okuldaki sınavlarda da böyledir. Doğru ölçme aracı geliştirirseniz ölçmek istediğiniz şeyi doğru ölçersiniz. Öğretmenin ihtiyacını belirleme de aslında doğru bir öğretmen değerlendirme sistemi ile olur. Sonucunda da çıkan eksikliği öğretmen de kabul eder ve eğitim almaya istekli olur. Dışardan gelen hiçbir şeyi kabul etmiyor öğretmen. Siz en başarılı eğitimi de sunsanız sıkılıyor ya da istemiyor. Gelişim doğru değerlendirme ile olur. Tespit edemediğimiz eksikliği nasıl geliştirelim. (E2 altı yıllık tecrübe)

Most of the teachers ($f=12$), principals ($f=5$), and experts ($f=5$) participants suggested that the effectiveness of the training should be determined through observations in the classroom and interviews with teachers. Participants stated that the efficacy of education could be determined best if the classroom instruction is observed or if one-to-one interviews can be conducted with the teacher. They also said it could not be

determined whether the teacher uses what she has learned in education without observing or interviewing. One of the teachers explained the need for assessing the effectiveness of the training through observations in the classroom as follows:

Okay, I took this training, but then did I practice what I learned? It is only evident through classroom observation. Now, for example, we have all studied mathematics literacy. I'm trying to teach math in class with this education. But I stayed behind other teachers in terms of subject matter. Now, in this case, they do not apply what they have learned in education, and nobody knows it. (T6 six years of experience)

Tamam, ben bu eğitimi aldım ama sonrasında öğrendiğimi uyguladım mı uygulamadım mı? Bu ancak sınıf gözlemi ile belli olur. Şimdi mesela hepimiz matematik okuryazarlığı eğitimi aldık. Ben sınıfta matematiğim bu aldığım eğitimle öğretmeye çalışıyorum. Ama bir baktım diğer öğretmenlerin gerisineyim konu olarak. Şimdi bu durumda onlar eğitimde öğrendiklerini uygulamıyor ve bunu kimse bilmiyor. (T6 altı yıllık tecrübe)

Some of the teachers (f=7), principals (f=4), and experts (f=4) also stated that professional development could be achieved by being aware of the good practices of other teachers in the classroom. For example, the school principal explained this suggestion:

I even noticed a very effective practice in a classroom, and when the lesson was over, I said, "This is a very different practice. It is very effective. Do you share it with other teachers?" She told me that sometimes she shared what she had done, and sometimes she forgot to share it. I'm sure another teacher would be very impressed. Therefore, I think it is essential for teachers to see the practices in each other's classrooms in terms of professional development. (P4 seventeen years of experience)

Bir sınıfta çok etkili bir uygulama fark ettim ve ders bitince dedim ki "bu çok farklı bir uygulama çok ta etkili. Diğer öğretmenlerle paylaşıyor musun?" o da bana yaptıklarını bazen paylaştığını bazen de paylaşmayı unuttuğunu söyledi. Benim yerimde başka bir sınıf öğretmeni olsaydı eminim ki o da çok etkilenecekti. Öğretmenlerin birbirinin sınıfındaki uygulamaları görmesi mesleki gelişim açısından bence çok önemli. (P4 on yedi yıllık tecrübe)

4.2 Second Part of The Study

In the second part of the study, the researcher developed an initial model in line with the opinions obtained from the first part of the study. Since this study aims to develop

a participatory teacher evaluation in this second participatory process initial model was introduced to the participants, and the opinions of the participants were taken with focus group interviews. In this section, the development process of the initial model and the codes and answers that emerged as a result of the focus group discussion are explained in detail.

4.2.1 The Initial Participatory Teacher Evaluation Model

While developing this initial model, the recommendations of the participants were considered as the basis. The purpose, the important points that should be included in the model, the qualification to be evaluated, the people who will make the evaluation, the qualification areas that these people will evaluate, evaluation tools, and the timeline are included in the initial model.

In line with the recommendations, the purpose of the initial model and the important issues proposed to be included were arranged and listed by the researcher. The purpose of the model was included in the model as "to evaluate teacher qualifications and to provide systematic support for teacher professional development that depends on the teacher evaluation process." On the other hand, the topics that the participants frequently mentioned about the important aspects of the model were listed under the name of important issues:

- Providing training for both evaluators and teachers to be evaluated
- Giving balanced feedback, including both strengths and weaknesses
- Ensuring confidentiality with face-to-face reporting
- Ensuring defensibility through face-to-face meetings
- Using multiple data sources for reliable information
- Making a holistic assessment at different times and by different people

In the first part of the study, the qualification areas that the participants found appropriate to be evaluated were determined as "planning and preparation," "instruction," "monitoring and managing learning," "communication and collaboration," "service to the school," and "professional development." In the focus

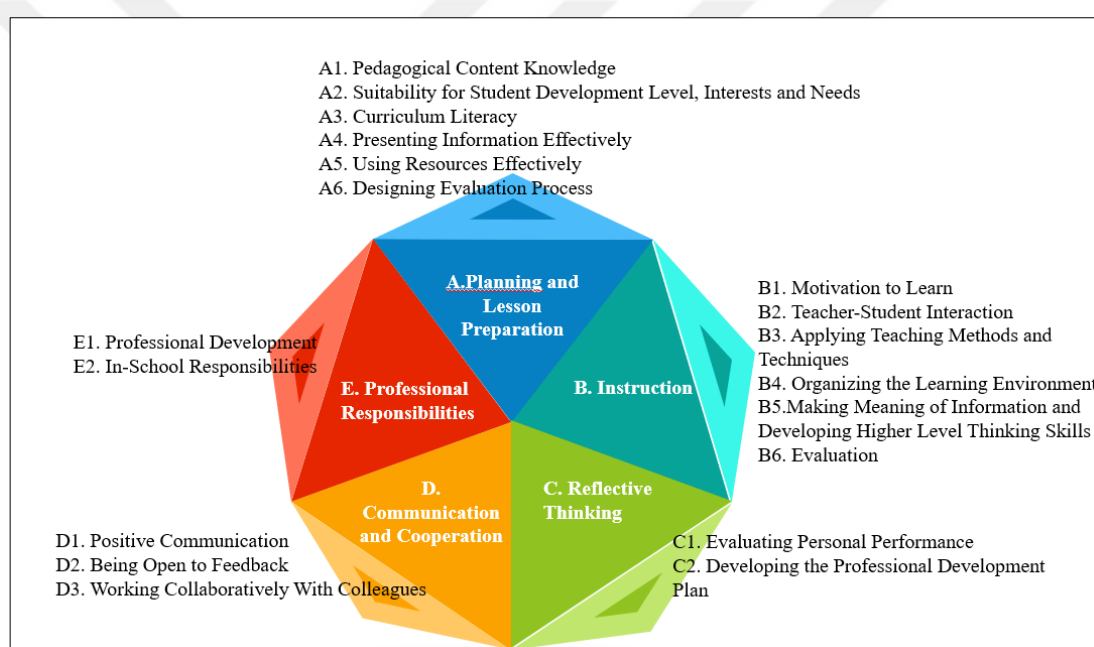
group meetings, it was planned to get opinions on these qualification areas and to provide the participants with the opportunity to change or add these areas when they deem necessary. For this opinion-taking process, sub-dimensions were added by the researcher to include the scope of the qualification and make the qualification area more understandable. While adding sub-dimensions, remarkable resources in the field of teacher evaluation (Bookhart, 2020; Danielson & McGreal, 2000; Danielson, 2007; Danielson, 2013; Graham et.al, 2015; Marzano & Toth, 2013; Marshall, 2013; Robinson & Aronica, 2016; Shulman, 1986; Stronge, 2018; Tucker ve Stronge, 2005) were considered. In addition, teacher qualification areas developed by the Ministry of National Education General Directorate of Teacher Training and Development were also examined (MoNE, 2017). With the help of these resources examined, the sub-areas covered by the "Instruction" qualification area are determined as "Motivation to Learn, Teacher-Student Interaction, Application of Teaching Methods and Techniques, Organizing the Learning Environment, Making Meaning of Information and Developing Higher Level Thinking Skills, Evaluation". The sub-areas covered by the "Communication and Collaboration" qualification area are determined as "Positive Communication" and "Being Open to Feedback from Colleagues".

Some qualification areas and sub-areas were developed with both the opinions of the participants and the relevant literature research. For instance, "Monitoring the Learning" was suggested by the participants, but this qualification was not given as a separate title since it took place as sub-fields under the qualification areas of both planning and instruction. In the first part of the study participants explained "service to school" qualification as professional and in-school responsibilities of teachers such as making contributions to the development of the school, participating in the activities carried out in the school, keeping duties in the school, and taking responsibility in ceremonies. "In-School Responsibilities" has been determined as a sub-dimension for this qualification field. On the other hand, in some sources in the literature, "professional development" is not given as a separate heading, since professional development is also included under this qualification area, and it is specified as a sub-dimension.

In the first part, the participants mentioned the importance of the teacher's self-expression and self-evaluation, but they did not define these important statements as a field of qualification. Although the participants did not suggest a qualification area, it was decided to add another qualification under the title of "Reflective Thinking" in line with the literature review. When both the opinions of the participants and the sources in the literature are examined, qualification areas and sub-areas that should be included in the initial model are determined as in Figure 4.1.

Figure 4.1

Teacher Qualification Areas and Subdimensions Included in The Initial Model



A schedule was proposed to the participants, considering the suggestions of the participants about the evaluators, the methods to be used in the data collection process, and the time of the evaluations during an academic year. In the first part of the study, participants suggested observation, debriefing before and after the classroom observations, and self-evaluation to evaluate the instruction. They did not make any suggestions regarding data collection tools to be used in assessing other qualification areas (communication and cooperation, professional responsibilities). Data collection tools suggested by the researcher for these areas are also included in the schedule. The overall design of the initial model is included in Appendix G.

4.2.2 The Opinions of The Participants Regarding the Initial Participatory Teacher Evaluation Model (Research Question 5)

This initial model was included in the introductory part of the focus group interview form, and the researcher consulted two professors on the thesis committee for the content and face validity of the initial model and focus group questions. Focus group meetings were organized, and teachers, principals, and experts participated in four different focus group meetings. The initial model proposed with the opinions of the participants and the additions made by the researcher through reviewing the literature was introduced to the teachers, experts, and principals in each focus group. In these interviews, opinions were received about each element of the proposed model and the evaluation process. In addition, in the first part, it was not clear on some issues, such as feasibility and whether the lesson plans would be used within the scope of this model. Opinions on these issues were also received in the focus group meetings.

4.2.2.1 Important Issues Planned to be Included in the model

The opinions and suggestions that emerged from the focus group discussions on the important issues in the initial model were analyzed. The codes and themes that emerged as a result of this analysis are given in Table 4.19.

Table 6

Codes and Themes for The Important Issues in The Initial Model

Theme 1. Evaluation Process
<i>Important Issues</i>
Training for evaluators and teachers being evaluated *
Repetition of training
Multiple data sources*
Suggestions for providing consistency and objectivity
Practical training
Classroom teachers working at different grade levels
Simultaneous observation by two different observers

* This code was also revealed in the first part of the research.

All the participants agreed that both the evaluator and the teacher to be evaluated should be trained before implementing the evaluation model. Some of the teachers (T5, T8, T9, T11, T13, T20, T21, T29, T31, T32) also mentioned that to trust the model; they should be trained to be familiar with the purpose and the process of the evaluation model. One of the teachers explained the importance of training:

The school principal made class observations this year. I don't know why this observation was made and what happened afterward. All of this should be explained to us from the beginning of the training. Or for example, what I should do in the classroom, how I should behave during the observation, and what I have to say to my students should be clearly stated in training. Otherwise, what we're facing is total chaos in the classroom. (T5 seventeen years of experience)

Sınıf gözlemleri bu yıl okul müdürü tarafından yapıldı. Bu gözlemin neden yapıldığını ve sonrasında ne olduğunu bilmiyorum. Bütün bunlar bize eğitimlerin başından itibaren açıklanmalı. Veya örneğin, sınıfta ne yapmalıyım, gözlem sırasında nasıl davranmalıyım ve öğrencilerime söylemem gerekenler eğitimlerde açıkça belirtilmelidir. Aksi takdirde, karşı karşıya olduğumuz şey sınıfta tam bir kaos. (T5 on yedi yıllık deneyim)

Some participants (T11, T13, T15, T27, E3, E6, P1, P2) mentioned that these training should be held during the semester and renewed periodically. One of the experts (E6 *five years of experience*) explained the importance of repetition of the training as; "*Sometimes what is said at the beginning of the academic year can be forgotten at the end of the term. I think it is important to repeat this training, that is, to make occasional reminders. (Bazen akademik yılın başında söylenenler dönem sonunda unutulabilir, bu yüzden bu eğitimleri tekrarlamak, yani ara sıra hatırlatmak yapmak önemlidir.)*"

Using multiple data sources to provide reliable information and holistic assessment done by different people at different times were considered one of the essential features of that model. Some of the participants (T3, T10, T15, T27, T29, T30, T32, E6, P5, P2) stated that the teacher should trust the evaluation results and that a single measurement would not be enough for them to trust. On the other hand, some of the participants (T9, T13, T11, T21, T15, T22, T24, T26, T27, T29, T30, T32, P1, E1, E3) stated that conducting the assessment continuously by different people for one year would provide more accurate and reliable information about teachers' weaknesses.

One of the principals stated:

It will not be enough to observe a single lesson. The most important thing that should be in an evaluation model is to make more than one observation made by different people. While everything is perfectly designed in an evaluation model, a single observation will not provide enough information to anyone. (P1 thirteen years of experience)

Tek bir dersin gözlemlenmesi yeterli olmayacaktır. Bir değerlendirme modelinde olması gereken en önemli hususun, farklı insanlar tarafından yapılan birden fazla gözlem yapmak olduğunu düşünüyorum. Her şey bir değerlendirme modelinde mükemmel bir şekilde tasarlanmış olsa da tek bir gözlem kimseye yeterli bilgi vermeyecektir. (P1 on üç yıllık tecrübe)

One teacher explained:

Perhaps the most important and always complaining is an evaluation with a single data source or a single measurement. For example, I am not the same in every lesson. Many times, when the lesson is over, I say to myself that I should repeat this lesson that was not well understood or was excellent. I think I am average in all these courses. (T29 eight years of experience)

Belki de en önemlisi ve her zaman şikâyet ettiğimiz şey tek bir kaynak veya tek bir ölçüm ile değerlendirme yapılması. Her derste ben aynı değilim mesela. Pek çok kez ders bittiğinde kendime diyorum ki bu dersi tekrarlamalıyım iyi anlaşılmadı veya çok iyiydi diye. Bence ben aslında tüm bu derslerin ortalamasıyım. (T29 sekiz yıllık tecrübe)

In the focus group interviews, the importance of ensuring consistency between the observers and the observation results was mentioned when more than one observation was made. Some of the participants (T3, T5, T8, T26, T28, T29, T30, T32, P3, P5, E1, E6) emphasized that it was important for the teacher's behavior to show continuity in order to make the right decision about the teacher. According to these participants, they emphasized that it would be appropriate to consider the results of observations only if this behavior was constantly repeated and if it was revealed in the same way by different observers, that is, if it was consistent. Some participants (T11, T8, T15, T29, T30, T31, P3, P5, E6, E7) explained that this consistency or inconsistency could be caused by the observer. To prevent this inconsistency, it is important for the evaluators to receive practical training via evaluating real classroom environments or scenarios. Two different teachers expressed the need for consistency as follows.

If observers use the observation form differently, there is an inconsistency here. Everyone should understand the form in the same

Gözlemciler gözlem formunu farklı kullanıyorsa burada bir tutarsızlık olur. Herkesin formu aynı şekilde anlaması lazım. Gözlemciler bu

way. After all, the observers will be the teachers and principals in this school. These observers need to practice using the form, just like piloting. These applications can also be in training, for example. (T8 eighteen years of experience)

okuldaki öğretmenler ve yöneticiler olacak sonuçta. Bu gözlemcilerin formu kullanarak uygulamalar yapması lazım aynı pilot yapar gibi. Bu uygulamalar eğitimlerde de olabilir mesela. (T8 onsekiz yıllık tecrübe)

One observer came, observed, and said that the classroom management is weak; the other one observed another lesson and said that the classroom management is very good. What about now? Is my classroom management good or bad? It may be like this; for example, four observations were made during the period, and classroom management is weak in all three of them. Then this result should, of course, be considered. (T32 eight years of experience)

Bir gözlemci geldi izledi ve dedi ki sınıf yönetimi zayıf, diğeri başka dersi gözlemledi sınıf yönetimi çok iyi dedi. Ne olacak şimdi, benim sınıf yönetimim iyi mi kötü mü? Şöyle olabilir ama diğer gözlemlerde de bu böyle çıkıyor mesela dönemde dört gözlem yapılmış üçünde de sınıf yönetimi zayıf. O zaman bu sonuç dikkate alınmalı tabiki. (T32 sekiz yıllık tecrübe)

A teacher explained her opinions about the observation of two people at the same time:

For example, if the head of the department and a classroom teacher enter the class together, they can speak to correct misunderstandings after observation. Or two different classroom teachers can enter at the same time. They can clarify points that are invisible to each other and points that are not clear. Or they both fill two separate observation forms, but then the consistency between these forms can be checked. Then it turns out that one teacher evaluated the other objectively. (T24 five years of experience)

Örneğin, zümre başkanı ve bir sınıf öğretmeni sınıfa birlikte girerse, gözlemden sonra yanlış anlamaları düzeltmek için konuşabilirler. Ya da iki farklı sınıf öğretmeni de aynı anda girebilir. Birbirlerinin gözünden kaçan noktaları ve net olmayan noktaları netleştirebilirler. Ya da ikisi de ayrı iki gözlem formu doldurur ama sonra bu formlar arasındaki tutarlılığa bakılabilir. O zaman bir öğretmen diğerini objektif değerlendirdi mi çıkar ortaya. (T24 beş yıllık tecrübe)

The participants were asked what could be done to ensure objectivity. Most of the teachers stated that the teacher who will take the course could be selected from the upper or lower classes to ensure objectivity. A teacher stated:

For example, a lower-class teacher can go to an upper class for observation. Because close friendship between teachers at the same grade level can prevent being neutral. For example, a 4th-grade teacher can observe 1st-grade teachers. (T25 eight years of experience)

Örneğin, bir alt sınıf öğretmeni gözlem için bir üst sınıfa gidebilir. Çünkü aynı sınıf düzeyinde öğretmenler arasındaki yakın dostluk tarafsız olmayı engelleyebilir. Örneğin, 4. sınıf öğretmeni 1. sınıf öğretmenlerini gözlemleyebilir. (T25 sekiz yıllık tecrübe)

Another teacher mentioned:

In other schools, teachers teach at a single grade level. For example, teachers constantly teach 1st grade. But here, as classroom teachers, we teach all four grade levels, and we have the right to teach at every grade level. So, I would have no problem, for example, if the first-grade teacher observed me. (T11 twenty years of experience)

Diğer okullarda öğretmenler tek bir sınıf düzeyde ders vermektedir. Örneğin öğretmenler sürekli 1. sınıfta okutuyorlar. Ama biz burada sınıf öğretmenleri olarak dört sınıf düzeyinin hepsini öğretiyoruz ve her sınıf seviyesinde öğretme hakkımız var. Bu nedenle, örneğin beni birinci sınıf öğretmeni gözlemlemese hiç sorun yaşamazdım. (T11 yirmi yıllık tecrübe)

Participants were also asked if it would be a solution to provide objectivity if more than one person simultaneously participated in the observation process. One of the experts (E3 ten years of experience) explained as “In fact, two teachers can enter the classroom at the same time. After that, we can also look at inter-observer consistency after observation. In this way, many problems in achieving objectivity are solved. (Aslında, iki öğretmen aynı anda sınıfa girebilir. Bundan sonra, gözlemden sonra gözlemciler arası tutarlılığa da bakabiliriz. Bu şekilde objektiflik sağlamak için birçok sorun çözülür.)”

Some teachers (T8, T9, T10, T11, T15, T23, T24) stated that they were worried that the classroom observations could cause differences in their and their students' behaviors, and in this case, what happened during the observation might not reflect the natural environment of the classroom. When these concerns were raised in the focus group interviews, they stated that a certain familiarization period was needed to overcome this situation. The observation process would be accepted as a normal process after a while. One of the teachers explained this concern and the solution:

I agree that my student's behavior will be different, especially in the first observations. For example, I have a student who speaks a lot and always asks questions; maybe he will hesitate and not talk. Even I can be very excited. It makes me uneasy, but perhaps we will get used to being comfortable as we talked here over time, and students won't care too much. (T23 four years of experience)

Ben de katılıyorum özellikle ilk gözlemlerde öğrencilerimin davranışı farklı olacaktır. Örneğin, çok konuşan ve her zaman soru soran bir öğrencim var, belki tereddüt edecek ve konuşmayacak. Ben bile çok heyecanlanabiliyorum. Bu beni huzursuz ediyor ama burada konuştuğumuz gibi rahat olmaya belki de zamanla alışacağız ve öğrenciler de çok fazla umursamayacaklar. (T23 dört yıllık tecrübe)

Another teacher explained:

For example, in my old school, observation had become a very normal process. Students were observers in their class from an early age. The students were very used to it. If someone came to class, nobody would care; it was normal for them now. The door of the classroom was always open. I remember even not realizing that the observer was there. Here too, it takes time, but it can be like this. (T9 three years of experience)

Örneğin, eski okulumda gözlem çok normal bir süreç haline gelmişti. Öğrencilerin küçük yaşlardan itibaren sınıfına gözlemci geliyordu. Öğrenciler buna çok alışkınlardı. Birisi sınıfa gelirse kimse umursamazdı artık onlar için bu durum çok normaldi. Sınıfın kapısı her zaman açıktı. Gözlemcinin orada olduğunu fark etmediğimi bile hatırlıyorum. Burada da zaman lazım ama böyle olabilir. (T9 üç yıllık tecrübe)

4.2.2.2 Organization of the Qualifications

The opinions and suggestions that emerged from the focus group discussions on the teacher qualifications to be evaluated in the initial model were analyzed. The codes and themes that emerged as a result of this analysis are given in Table 4.20.

Table 4.20

Codes and Themes for The Qualifications to be Evaluated in The Initial Model

<i>Theme 1. Evaluation Process</i>
<i>Organization of the Qualifications</i>
Reorganizing or changing qualifications
Equal importance to each qualification
Same qualifications for experienced and novice teachers

* This code was also revealed in the first part of the research.

In all four focus group meetings, most of the participants emphasized that the qualifications were too many; some of them were intertwined and included each other. Therefore, with some of the participants (T8, T11, T13, T25, T28, T31, P1, P5, E1, E3, E6, E7), it was decided to reorganize the sub-dimensions in the planning and preparation qualification area. For example, it was decided to include all sub-dimensions of "Relevance to Student Development, Interests and Needs," "Providing information effectively," and "Effective use of resources" under the dimension of "designing the instruction." While this reorganization was being made, the opinions of all the other participants were taken, and the final version of the qualification area was formed with the approval of everyone.

It has been stated that the "Classroom Organization and Management" qualification, which is a vital area, especially for classroom teachers, should be included as a separate area and should take place instead of "Organizing the Learning Environment." Therefore, with the contribution of most of the participants, it was decided that "Classroom Organization and Management" should take place instead of "Organizing the Learning Environment." After this arrangement, the approval of the other participants was obtained, and it was stated that it would be more understandable as it is.

Some participants stated that the qualification area of "Making Meaning of Information and Developing Higher Level Thinking Skills" was not understood. Therefore, this title should be changed and said more clearly. For this purpose, this area was organized as "Making sense of knowledge and development of skills," and it is divided into two sub-areas: "Ensuring Student Interaction with New Knowledge" and "Helping Students to Apply and Deepen New Knowledge." While some participants (T3, T9, T11, T10, T15, T22, T24, T32, P1, P3, P5, E1, E3, E6, E7) actively participated in the creation of these areas, other participants expressed their approval.

The participants that the two sub-dimensions in the reflective thinking qualification area, "Evaluating personal performance" and "Creating a professional development

plan," did not fully meet this qualification area. Furthermore, it was decided that the creation of a professional development plan was not suitable for reflective thinking because this plan would be created with the cooperation of the principal and the teacher at the end of the evaluation. Therefore, this qualification area has been changed to "Reflecting on Teaching and Evaluating Own Performance" with the active participation of some participants (T5, T10, T11, T13, T15, T22, T24, T26, T27, T30, T32, P1, P2, P5, E1, E3, E6, E7) and other participants expressed their approval.

It was stated that the field of "being open to feedback from colleagues" in the field of communication and cooperation already exists in the other two areas, "Positive communication" and "Working in Collaboration with Colleagues." Therefore, with the active participation of some participants (T9, T10, T11, T13, T15, T21, T24, T30, T32, P1, P2, P5, E1, E3, E6, E7), this area was removed, and other participants expressed their approval.

In general, the participants stated that the teacher qualifications in the model and the expressions mentioned below covered the areas they expected to be in. In line with this requirement, participants noted that adding or subtracting items or expressions from these qualifications was unnecessary. For example, one teacher (T11 *Twenty years of experience*) explained this situation as *"It is essential to understand the outcomes in the program correctly. I thought that there should be a competency in understanding the outcomes. Still, then I realized that this is already included in other competencies. (Programdaki kazanımları doğru anlamak çok önemli. Kazanımları anlamak ile ilgili bir yeterlik olmalı diye düşündüm ama sonra farkettimki bu diğer yeterlik alanlarının içinde yer alır zaten.)"*

Another teacher explained this situation as:

It is essential to get feedback from colleagues about communication. I think it should be not only to get feedback from colleagues but also to give feedback to colleagues, but in line with our discussion here, I have decided that this competence is already in communication and

İletişim ile ilgili meslektaşlardan geri bildirim almak çok önemlidir. Sadece meslektaşlardan geri bildirim almanın değil, aynı zamanda meslektaşlara geri bildirimde bulunmanın da olması gerektiğini düşünüyorum, ancak buradaki tartışmamız doğrultusunda, bu yeterliliğin zaten iletişim ve iş birliği

cooperation. (T21 three years of experience) içinde olduğuna karar verdim. (T21 üç yıllık tecrübe)

Participants were asked whether the qualifications planned to be in the model should be different for novice and experienced teachers. In the light of the answers received, it was revealed that all the qualifications mentioned should be the same for both novice and experienced teachers. Some participants (T3, T9, T13, T15, T23, T26, T28, T29, T30, P1, P3, P5, E3, E6) stated that all teachers should have all the qualifications mentioned in the model. In contrast, others (T11, T27, T31, T32, E7) noted that the number of new teachers in this school was low. A few participants (T3, T8, T13, P2, P3) stated that many criteria were considered when recruiting teachers to the school and that it was expected to have specific qualifications regardless of whether the teacher was new or experienced. One of the teachers (T9 three years of experience) explained this situation, “I think that the assessment criteria of both novice and experienced teachers should be the same. I don’t think the experienced teacher is better than the new teacher. I think the new teacher is more knowledgeable in the field. (Hem yeni hem de deneyimli öğretmenlerin değerlendirme kriterlerinin aynı olması gerektiğini düşünüyorum. Tecrübeli öğretmenin yeni öğretmenden daha iyi olduğunu düşünmüyorum. Bence yeni öğretmen bu alanda daha bilgili.)”

One teacher explained this situation:

I think it is not a problem to evaluate inexperienced and experienced teachers within the same qualification areas. In this school, everyone comes with a specific experience [...] The data collection tools to be used here must be a scoring key. If we state how experienced we are at the beginning of this form, I think it is understandable whether the deficiencies arising from observation are due to years of experience. (T32 two years of experience) Tecrübesiz ve deneyimli öğretmenlerin aynı yeterlik alanları içinde değerlendirilmesinin sorun olmadığını düşünüyorum. Bu okulda herkes belli bir deneyimle geliyor... Burada kullanılacak veri toplama araçlarının bir puanlama anahtarının olması çok önemlidir. Ne kadar tecrübeli olduğumuzu bu formun başlangıcında belirtirsek, gözlemden kaynaklanan eksikliklerin tecrübe yılı nedeniyle olup olmadığı anlaşılır bence. (T32 iki yıllık tecrübe)

One principal explained this situation:

The new teacher should be better equipped with in-field knowledge, but the experienced teacher should keep the field knowledge up to date. The same is true for classroom teaching, so every teacher should be very good at these competencies. For example, when we recruit teachers at our school, we expect classroom management to be very good for a new and experienced teacher. (P2 twenty-two years of experience)

Yeni öğretmen alan bilgisi anlamında daha donanımlı olmalıdır, ancak deneyimli öğretmen alan bilgisini güncel tutmalıdır. Aynı şey sınıfta öğretim için de geçerlidir, bu yüzden her öğretmenin tüm bu yeterliliklerde çok iyi olmalıdır. Mesela, okulumuzda öğretmen işe alımı yaparken sınıf yönetiminin yeni ve tecrübeli öğretmen de olasa çok iyi olmasını bekliyoruz. (P2 yirmi iki yıllık deneyim)

When the participants were asked whether there was a priority or importance order among these qualifications, they stated that each qualification was essential and that they should be together for the teacher evaluation to be holistic. Some participants (T5, T8, T10, T15, T22, T25, T26, P1, P2, E1, E3) stated that it was more important to determine the strengths and improvement aspects of each area rather than providing them with a total score and that having a specific score would reveal a competitive environment among the teachers and they named this as an undesirable situation. One of the teachers (T8 eighteen years of experience) explained the equal importance of the qualification "For example, a teacher has a good relationship with their student, but their communication with other teachers and principals is very bad. I think she's not a good teacher if he doesn't cooperate with other friends he works with. (Örneğin bir öğretmenin öğrencisiyle arası iyi ama diğer öğretmen ve müdürlerle iletişimi çok kötü. Çalıştığı diğer arkadaşlarıyla işbirliği yapmıyorsa bence iyi bir öğretmen değil.)"

Another (T26) teacher expressed her views on this subject as follows "If a teacher does not do in-school responsibilities, this causes injustice. This situation causes a lot of tension in the school. Therefore, it is crucial and cannot be less important than others. (Örneğin bir öğretmen okul içi sorumluluklarını yerine getirmezse bu adaletsizliğe neden olur. Bu durum okulda çok fazla gerginliğe neden olur. Bu nedenle bu alanda çok önemlidir ve diğerlerinden ayrı ya da az tutulamaz.)"

One expert explained this situation:

The most important thing to mention here is whether there will be scoring in this model or it will be enough to identify only strengths and weaknesses in the end. Secondly, a development plan can be prepared for each area, but it will be challenging to create a development plan if a score is given. Therefore, whether the learning areas are all equally important or which are most important depends very much on our purpose. In such a model, each can be of equal importance, for example (E3 ten years of experience)

Burada bahsetmemiz gereken en önemli şey, bu modelde bir puanlamanın olup olmayacağı veya sonunda sadece güçlü ve zayıf yönlerin belirlenmesinin yeterli olup olmayacağıdır. İkincisi ise, her alan için bir kalkınma planı hazırlanabilir, ancak puan verilirse bir kalkınma planı oluşturmak zor olacaktır. Dolayısıyla öğrenme alanlarının hepsinin eşit derecede mi önemli olduğu ya da hangisinin en önemli olduğu amacımıza çok bağlıdır. Böyle bir modelde her bir eşit önemde olabilir mesela. (E3 on yıllık tecrübe)

One principal explained this situation:

It does not seem right to give me points, so there is no clear order of importance. When you earn points, we can say that the teacher has passed or couldn't pass. However, it should be identified in which area there is a deficiency, and she got a low score from here or how we can improve her in this area. (P1 thirteen years of experience)

Bana puan vermek doğru gelmiyor, bu yüzden açık bir önem sırası yok. Puan kazandığınızda, sanki öğretmen geçti ya da kaldı diyebiliyoruz. Oysa hangi alanda ne eksiği vardı da buradan düşük puan aldı ya da ne yaparsak bu alanda onu geliştirebiliriz diye bakılmalı. (P1 on üç yıllık tecrübe)

4.2.2.3 Schedule of the Model

The opinions and suggestions that emerged from the focus group discussions on the schedule of the initial model were analyzed. The codes and themes that emerged as a result of this analysis are given in Table 4.21.

Table 4.21*Codes and Themes for the Schedule of The Initial Model*

<i>Theme 1. Evaluation Process</i>
<i>Schedule of The Initial Model</i>
Science and math teachers*
Classroom teacher evaluating communication and collaboration
Using lesson plans as a data source
Observations*
Debriefing before and after observation *
Self-evaluation*

* This code was also revealed in the first part of the research.

Participants have also explained details in the schedule like whom, how, and what specific time of the academic year evaluations will be done. Participants were asked to examine each column in the table where the planned schedule of the model was explained. When the participants were asked whether the qualifications could be evaluated by the people stated, they mentioned that as long as they are trained about the evaluation process and communicate positively, the people in the table would be appropriate to evaluate these qualifications. Regarding the schedule, the teachers stated that the number of observations in some months (for example, October and January) could be increased. One teacher explained the importance of evaluator training as follows.

<i>Of course, these people can evaluate, but I also question their competencies. So are they trained? Even if I am training to improve my teaching knowledge, those who evaluate me should receive regular training. (T25 eight years of experience)</i>	<i>Bu kişiler tabiki değerlendirebilir ama yeterliklerini de sorgularım. Yani eğitim almışlar mı? ben bile öğretmenlik bilgilerimi yenilemek için eğitim alıyorsam beni değerlendirende düzenli eğitim almalı (T25 sekiz yıllık tecrübe)</i>

One teacher explained this situation:

<i>I think the people in this table can evaluate us. I think what matters is whether the observer knows how to communicate. Observers should</i>	<i>Bence bu tablodaki kişiler tarafından değerlendirilebiliriz. Bence önemli olan, gözlemleyen kişinin nasıl iletişim kuracağını bilip bilmediğidir.</i>

not be superior or advised. I must believe that feedback is for my development. The observer should treat us just like a mirror. Oh, and training is essential. Both the evaluator and the evaluated should receive training at regular intervals. (T31 *six years of experience*)

Gözlemciler üstünlük taslamamalı ya da tavsiye vermemelidir. Yani ben geri bildirimin gelişimim için olduğuna inanmalıyım. Gözlemci bize sadece bir ayna gibi davranmalıdır. Birde eğitim önemli. Değerlendiren de değerlendirilenlerde belli aralıklarla eğitim almalıdır. (T31 *altı yıllık tecrübe*)

Among the people mentioned, the most frequently discussed evaluators were science and math teachers assessing the lessons. While some teachers (T3, T10, T11, T24, T25, T26, T30, T32) thought that these teachers could not understand the level of their classes, it was concluded that if the criteria stated in the observation forms were clear and understandable, this would not be a problem. Most of the participants (T3, T5, T8, T10, T13, T15, T22, T23, T26, T27, T30, T32, E1, E3, P1, P3, P5) stated that math and science teachers' observations were needed to get advice and to enrich what they know in this field. In addition, the participants noted that the observation made by these people was essential to provide facilities to the students by eliminating the deficiencies that may arise in mathematics and science, especially during the transition from 4th grade to 5th grade. One of the teachers (T15 *four years of experience*) explained, "If the math or science teacher is able to adapt to the level, then the observation will be meaningful. It should be possible that they evaluate the level according to the criteria we give them. (Matematik veya fen bilgisi öğretmeni seviyeye uyum sağlayabilirse gözlem anlamlı olacaktır. Bu ancak bizim belirlediğimiz kriterlere göre seviyeyi değerlendirmesi ile mümkün olur.)" Another teacher stated:

In other words, since these people will come from middle school, they should not look at students' behavior or classroom management. However, especially in the 4th grade, they should identify the parts we taught wrong or missing in science and mathematics lessons and give us feedback. (T13 *twelve years of experience*)

Başka bir deyişle, bu insanlar ortaokuldan gelecekleri için sınıftaki öğrencilerin davranışlarına veya sınıf yönetimine bakmamalıdır. Ancak özellikle 4. sınıfta, fen ve matematik derslerinde yanlış veya eksik öğrettiğimizi yerleri belirlemeli ve bize geri bildirimde bulunmalıdır. (T13 *on iki yıllık tecrübe*)

A teacher explained the importance of these observations:

Most straightforwardly, while teaching mathematics in middle school is algorithmic, we teach students to feel math in primary school and want them to understand mathematics. At the very least, the observer can note this. Students who move from 4th grade to 5th grade face great difficulty because an abstract education begins suddenly. They do not understand what they have learned, especially in the first months. For this reason, I find the observations of people from different branches very meaningful. (T23 four years of experience)

En basit şekilde, ortaokulda matematik öğretimi algoritmik iken, ilkokulda öğrencilere matematiği hissetmelerini öğretiriz ve matematiği anlamalarını istiyoruz. En azından gözlemci bunu fark keder. 4. sınıftan 5. sınıfa geçen öğrenciler büyük zorluk çekiyorlar çünkü birden soyut bir öğretim başlıyor. Özellikle ilk aylarda öğrendiklerini anlamıyorlar. Bu nedenle farklı branşlardan insanların gözlemlerini çok anlamlı buluyorum. (T23 dört yıllık tecrübe)

Most of the participants also stated that they are mostly communicating and working collaboratively with other classroom teachers so that this qualification could be evaluated not only by the head of the department but also this qualification should be evaluated by other classroom teachers. A teacher stated:

Evaluation of positive communication requires a process. We are always in contact with other teachers working in the same corridor, and we always talk about the teaching we do. I also find communication and collaboration beneficial in terms of professional development. Therefore, I want my communication to be evaluated by another class teacher. (T26 five years of experience)

Olumlu iletişimin değerlendirilmesi bir süreç gerektirir. Biz aynı koridorda görev yapan diğer öğretmenlerle hep temas halindeyiz ve yaptığımız öğretimle ilgili hep konuşuyoruz. İletişim ve iş birliğini mesleki gelişim açısından da çok yararlı buluyorum. Bu nedenle iletişimimin başka bir sınıf öğretmeni tarafından değerlendirilmesini istiyorum. (T26 beş yıllık tecrübe)

When the previous interviews were examined, it was determined that most of the teachers stated that the evaluation of the lesson plans should not take place because it would not reflect the process totally. With the help of the focus group interview, the participants were asked what they thought about using the lesson plan within such a model applied in line with the schedule stated in the table. The participants stated that there was no problem in evaluating the lesson plans as long as they were not seen as a

single resource and if the stages specified in this model were realized. One of the teachers explained this:

If the lesson plans in this model will not be determinant alone, evaluating the plans can be used because the classroom teaching is also examined. Sometimes lesson plans cannot be applied in the classroom, and we can involve other things according to the situation of the class at the moment. For example, if an interview will be held before the observation in this model and whether the prepared plan can be applied or not can be asked. I think it would be helpful to have the opportunity to explain this. (T27 ten years of experience)

Bu modeldeki ders planları tek başına belirleyici olmayacaksa planları değerlendirmek de sakınca yok çünkü sınıftaki öğretime de bakılıyor. Bazen sınıfta ders planını iç uygulanamaz ve o anda sınıfın durumuna göre başka şeyler de uygularız. Örneğin, bu modeldeki gözlemden önce bir görüşme yapılacaksa ve hazırlanan planın uygulanıp uygulanamayacağı da sorulabilir. Bence açıklama fırsatına sahip olmak faydalı olacaktır. (T27 on yıllık deneyim)

One of the principals explained:

I think the plan can be used if it contains specific criteria and standards. One of the criteria should be whether it addresses differences in the classroom. In the first meeting, I said that the planning logic here is not appropriate, but if the plan preparation logic is to be changed, if the plan will be observed in the lesson and teachers will not be expected to adhere to it one-on-one, it can be used in the assessment. (P3 sixteen years of experience)

Planın belli ölçütleri ve standartları içeriyorsa kullanılabileceğini düşünüyorum. Ölçütlerden biri, sınıftaki farklılıkları ele alıp almadığı olmalıdır. İlk görüşmede, buradaki plan mantığının uygun olmadığını söyledim, ancak plan hazırlama mantığı değiştirilecekse, plan derste gözlemlenecek ise ve öğretmenin birebir bağlı kalması beklenmeyecek ise değerlendirmede kullanılabilir. (P3 on altı yıllık tecrübe)

The participants were asked about their opinions on the appropriateness of the methods and data collection tools presented in the schedule. The participants stated that the tools used were appropriate and that the observation forms to be used should be well structured. Participants also noted that the form played a critical role in ensuring objectivity. In addition, they pointed out that the interviews conducted before and after the observation process were essential in ensuring objectivity. A teacher stated:

I think that the well-prepared form and the objectivity of the observation are directly related. Gözlemin objektifliği ile iyi hazırlanmış formun doğrudan ilişkili olduğunu düşünüyorum. Kesinlikle,

Indeed, the criteria in the form should be determined very well, and people to be observed with the form should be consulted. The clearer the items, the more truth will be reflected in the observed behavior. (T26 five years of experience)

formdaki kriterler çok iyi belirlenmeli ve form hakkında tüm gözlemlenecek kişilere danışılmalıdır. Maddeler ne kadar net olursa, gözlemlenen davranışta o kadar fazla gerçeği yansıtacaktır. (T26 beş yıllık tecrübe)

One of the principals explained this importance:

In my opinion, the person who comes to the classroom and observers should know about the class and the lesson before observing the process, whether it is a mathematics teacher, a science teacher, or a classroom teacher. It can be done by conducting interviews before observation. They don't have to spend a long time; they can collect all the necessary information in a short time. (P3 sixteen years of experience)

Bence, sınıfa gelen ve gözlemleyen kişi ister matematik öğretmeni, fen bilgisi öğretmeni veya sınıf öğretmeni olsun süreci gözlemlemeden önce sınıf ve ders hakkında bilgi sahibi olmalıdır. Bu, gözlemden önce görüşmeler yapılarak yapılabilir. Uzun olması gerekmez, kısa sürede de gerekli tüm bilgileri toplayabilirler. (P3 on altı yıllık deneyim)

One of the experts explained this importance as:

In fact, the form used here is significant. Before conducting the observation with the form, the observer can ask for the information she plans to explain in the process to be observed, the status of the class, and the characteristics of the students. You can also ask which criteria can be observed in this course. After the observation, the teacher should be asked how she assesses herself first. The observer should always ask the teacher for the notes she received in the observation, which really makes the correct assessment. (E3 ten years of experience)

Aslında, burada kullanılan form çok önemli. Formla birlikte gözleme girmeden önce öğretmene gözlemlenecek süreçte hangi bilgileri anlatmayı planladığını, sınıfın durumunu ve öğrencilerin özelliklerini sorabilir. Bu derste hangi ölçütlerin gözlemleyebileceği de sorulabilir. Gözlemden sonra öğretmene önce kendini nasıl değerlendirdiği sorulmalıdır. Gözlemci her zaman öğretmene gözlemde aldığı notlarını sormalıdır ki bu gerçekten doğru değerlendirme yapmayı sağlar. (E3 on yıllık deneyim)

4.2.2.4 Access to Evaluation Information and Providing Propriety

The opinions and suggestions that emerged from the focus group discussions on the process of accessing evaluation information and providing propriety for the initial model were analyzed. The codes and themes that emerged as a result of this analysis are given in Table 4.22.

Table 4.22

Codes and Themes for Accessing Evaluation Information and Providing Propriety

<i>Theme 2. Use of Evaluation Results</i>
<i>Access to Evaluation Information</i>
Importance of confidentiality*
Individual feedback (face to face) *
Teacher Development Plan (Functional Reporting)
Open communication*
<i>Propriety</i>
Balanced Evaluation*
No scores

* This code was also revealed in the first part of the research.

Participants agreed that access to evaluation information should be face to face to provide confidentiality and ensure defensibility. Most of the participants (T5, T9, T11, T13, T15, T21, T22, T24, T26, T27, T29, T30, T32, E3, E6, P1, P5, P2) mentioned that it was crucial to provide feedback or report by face-to-face interviews to correct misunderstandings and express themselves more clearly.

One teacher stated this as:

All teachers know the importance of always giving individual feedback to students. It is said everywhere to give your students personal feedback. It is the same for teachers, and it is essential to provide unique and face-to-face feedback to teachers. Also, I don't want my friends to hear what she says about me, or I don't want anyone to listen

Öğretmenlerin hepsi öğrencilere her zaman bireysel geri bildirim vermenin önemini bilir. Bu her yerde de söylenir öğrencilerinize bireysel dönüt verin diye. Öğretmenler için de aynıdır ve öğretmenlere bireysel ve yüz yüze geribildirim vermek çok önemlidir. Ayrıca ben hiçbir arkadaşımın benim hakkımda ne söylediğini duymasını istemiyorum ya da kimsenin benim

to my opinions. This feedback fikirlerimi de duymasını istemiyorum.
session must be between the Bu geri bildirim oturumu ben ve müdür
principal and me. (T30 eight years of arasında olmalıdır. (T30 sekiz yıllık tecrübe)
experience)

One of the principals (P2 twenty-two years of experience) expressed her ideas, “I think that only the feedback given in writing is not understood. A face-to-face expression gives both sides the opportunity to explain themselves more easily. (Ben sadece yazılı olarak verilen geri bildirimin anlaşılmadığını düşünüyorum. Yüz yüze ifade, her iki tarafa da kendilerini daha kolay açıklama fırsatı veriyor.)” Also, all the participants agreed that a balanced evaluation by giving feedback on both strengths and weaknesses was vital to conduct the reporting process properly. One teacher stated this:

For example, a form was prepared, the items in this form were very well determined, and scoring was made. For example, scoring was determined from one to three. The aim here is to improve me to get one point, but I would also like to be appreciated for three points. (T10 four years of experience)

Örneğin bir form hazırlandı ve bu formdaki maddeler çok iyi belirlendi ve bir puanlama yapıldı. Örneğin puanlama bir ila üç arasında belirlendi. Buradaki amaç, bir puan aldığım yerde kendimi geliştirmek olmalı, ama üç puan aldığım yer ile ilgili de takdir edilmek isterim. (T10 dört yıllık tecrübe)

In this model, at the end of each term, the participants were brought together with the reports obtained from all evaluators. The participants were also informed that the school principal and teacher would come together at the end of each period to prepare a development plan with the help of the reports obtained from all evaluators. The teacher is planned to support their professional development in line with this development plan. Participants’ opinions about this reporting process and performance evaluation meeting were taken. Most of the participants mentioned that the model’s reporting and performance meetings would benefit their professional development. Some participants (T8, T9, T11, T13, T22, T26, T30, T32) stated that it is essential that the report was sent to them before going to these meetings and that they would also prepare a plan and participate in the meeting. Some of the participants (T3, T10, T11, T15, T22, T27, T30, P1, E1, E3) stated that in these meetings, the principal and the teacher should agree on mutual conversations and the views of both sides should be taken into consideration. One of the teachers (T3 fourteen years of experience) explained this

reporting process “This way of reporting is very important for professional development. It is necessary to talk about this report, especially through mutual discussion, and to develop a common understanding. (Bu raporlama şekli mesleki gelişim için çok önemli görünüyor. Bu raporun özellikle karşılıklı tartışma yoluyla konuşulması ve ortak bir anlayış geliştirilmesi önemli tabi.)”. Another teacher (T22 *five years of experience*) also added as “With the data obtained from these observation results, the teacher can form a development plan and gather with the manager. The two plans can be reviewed and mutually agreed upon.” “Bu gözlem sonuçlarından elde edilen verilerle öğretmen de bir gelişim planı oluşturabilir ve yönetici ile görüşebilir tabiki. İki plan gözden geçirilebilir ve karşılıklı olarak kabul edilebilir.”. One of the experts (E3 *ten years of experience*) also agreed with the idea of the teachers stating, “It is crucial to reach a common decision and to emphasize that development is always at the forefront of this decision. (Ortak bir karara varmak ve bu kararda her zaman gelişimin ön planda olduğunu vurgulamak çok önemli.)”

In addition, teachers were asked how to design a model that supports their professional development differently from this model. Some of the participants (T15, T24, T25, T26, T27, T30) stated that a model which aimed professional development should not be based on giving scores. Furthermore, some participants (T3, T8, T10, T15, P3, E1) stated that it was essential to take different measurements from as many people as possible, and some participants (T11, T14, T15, T31, T32, E6, P1, P3) noted that professional development was not for competing with other teachers but for each teacher’s individual development. They should be held responsible for their development.

One of the teachers stated:

I do not find the evaluations made like the exam are correct. I don't think these evaluations are for my professional development. For example, the exam was held and graded. Let's say I got 70 points, and I was successful in this evaluation. Or I got 20, and I failed. I think such a judgment has been made to find a

Sınav gibi yapılan değerlendirmeleri doğru bulmuyorum. Bu değerlendirmelerin mesleki gelişimim için olduğunu düşünmüyorum. Örneğin, sınav yapıldı ve notlandırıldı. Diyelim ki 70 puan aldım ve bu değerlendirmede başarılı oldum. Ya da 20 aldım ve başarısız oldum. Profesyonel gelişim için değil,

reason, not for professional development but to enforce the teacher. (T26 *five years of experience*)

öğretmene yaptırım uygulamak için gerekçe bulmak için böyle bir yargıya varıldığını düşünüyorum. (T26 *beş yıllık tecrübe*)

One of the experts stated:

If we are talking about development, there should be repeated measurements. At the end of these evaluations, if many people have said that a teacher improved herself in a specific area, the study has achieved its purpose. I think it is essential not to score this area but to determine the deficiency in this area correctly (E1 *eight years of experience*)

Gelişimden bahsediyorsak, tekrarlanan ölçümler olmalı. Bu değerlendirmelerin sonunda, eğer birçok insan bir öğretmenin belli bir alanda kendini geliştirdiğini söylediyse, bu çalışma amacına ulaşmış demektir. Bu alana puan vermek değil de bu alandaki eksikliği doğru belirleyebilmek bence önemli olan (E1 *sekiz yıllık tecrübe*)

One of the principals stated:

In order to be a truly developmental model, the result should allow every teacher to update their knowledge about the teaching profession, learn new things, or as simple look critically at him. Assessment should be specific to each teacher. (P3 *sixteen years of experience*)

Gerçekten gelişimsel bir model olması için, sonuç her öğretmenin öğretmenlik mesleği hakkındaki bilgilerini güncellemesine, yeni şeyler öğrenmesine ya da en basitinden kendisine eleştirel bakmasına olanak tanımalıdır. Değerlendirme her öğretmene özgü olmalıdır. (P3 *on altı yıllık tecrübe*)

4.2.2.5 Feasibility of the Model

The opinions and suggestions that emerged from the focus group discussions on the feasibility of the model were analyzed. The codes and themes that emerged as a result of this analysis are given in Table 4.23.

Table 4.23

Codes and Themes for the Feasibility of the Model

Theme 2. Use of Evaluation Results
Feasibility
Willingness to use the model
Reducing the duties and responsibilities of the head of the department

Participants were asked to comment on the feasibility of this model. They also asked whether they would like to be evaluated or conduct evaluations with this model. Participants indicated that the periods specified for the implementation of this model were appropriate. The participants stated that it was essential that the evaluation does not occur in the first month of school and at the time it closes. Most of the participants (T3, T8, T10, T11, T14, T15, T20, T21, T27, T31, T321, P3, E1, E7) also stated that their own opinions were included in this model and that a collaborative model was designed so that they felt valuable. They did not see any objection to applying this model to reflect their views. Some participants (T9, T10, T11, T14, T15, T21, T22, T27, T32, P1, E1, E7) also stated that everything was indicated precisely in the table, which would allow practical applications. One of the experts (E7) stated, "I think that anyone who has a say in the model will willingly implement this model. Some of the participants (T10, T9, T11, T22, T25, P2, E3, E7) expressed as it is crucial to reduce the course load of the head of the department to make these evaluations feasible.

One of the teachers explained as:

<i>Obviously, our opinion about the evaluations has never been asked. Nobody asked me, "how should we evaluate you." I do everything in cooperation. This evaluation, which was prepared by asking us for our opinion, is an example of excellent collaboration. (T15 seventeen years of experience)</i>	<i>Açıkçası şimdiye kadar değerlendirmeler ile ilgili bizim fikrimiz hiç sorulmadı. Kimse bana seni nasıl değerlendirelim demedi. Ben iş birliği içinde her şeyi yaparım. Bu şekilde bize sorularak fikrimiz alınarak hazırlanmış bu değerlendirme çok iyi bir iş birliği örneği. (T15 on yedi yıllık tecrübe)</i>

Another teacher explained as:

<i>I wish this model had been applied immediately in our school. In fact, I don't know if my thoughts will be used when creating the model. But I believe this model will work as it will meet the needs of most teachers. (T3 fourteen years of experience)</i>	<i>Keşke bu model okulumuzda hemen uygulansaydı. Aslında, model oluşturulurken düşüncelerimin kullanılıp kullanılmayacağını bilmiyorum. Ama çoğu öğretmenin ihtiyacını karşılayacağı için bu modelin çalışacağına inanıyorum. (T3 on dört yıllık deneyim)</i>

4.2.2.6 The Guide and Tools of the Model

The opinions and suggestions that emerged from the focus group discussions on the guide and the tools the model were analyzed. The codes and themes that emerged as a result of this analysis are given in Table 4.24.

Table 4.24

Codes and Themes for the Guide and The Tools of the Model

<i>Theme 3. Guide and Tools of the Model</i>
<i>Need for a detailed guide, including</i>
Importance of the teacher evaluation
The model development process
The purpose of the evaluation
Detailed explanation of the qualifications
Use of data collection tools
Consistency, objectivity for the process
Openness and confidentiality for the results
<i>Need for developed tools</i>
Piloting the tools

When the participants were asked questions about the different dimensions of the initial model, and especially when asked to evaluate its feasibility, they stated that there was a need for a guide in which this model was explained in more detail. The common view of the participants in all groups was to explain in detail the purpose of this model, the qualification areas included in it, the tools to be used, the timetable, how the process would be carried out, and how the teachers would be informed. In addition, the participants emphasized that a guidebook could be prepared for this purpose and that this booklet is important in terms of ensuring unity in applying this model. For this purpose, it was decided to prepare a guide booklet at the end of the meeting. Different participants described the need for this guide as follows.

Actually, I want to add something. Now, for example, we are asking, and you are explaining some things to us here, but what will happen if you are not in practice? I wish there was a document where they were written

Aslında bir şey eklemek istiyorum. Şimdi mesela biz soruyoruz ve sen bize buradaki bazı şeyleri açıklıyorsun ama uygulamada sen olmazsan ne olacak. Keşke bunların yazılı olduğu ve tek tek açıklandığı bir

and explained one by one. Then it may be easier for us to implement and understand. (T29 *eight years of experience*)

The qualifications are quite extensive, but what if everyone understands something different? They cannot make a correct assessment. There are still many teachers here who confuse these areas. An explanatory and written guide is needed. The criteria of the competencies should also be explained, that is, the indicators should be clear. (P3 *sixteen years of experience*)

Now you're telling me, I am trying to guess what was that? It's been years since I graduated. Wish there was a teacher's guidebook that explained the purpose, the process, and the result... There was something like this in the old version of the curricula, and it helped me a lot. (T3 *fourteen years of experience*).

Yes, this guide can even be used in training. Everyone can easily reach it. Both the evaluator and the evaluated person know what kind of model it is and apply it more willingly. There may be things we want to add or remove from this guide; even these can be asked to us. It is very important that the guide is understandable and useful. (T21_{three} *years of experience*).

belge olsa. O zaman bizim için uygulaması ve anlaması daha kolay olabilir. (T29 *sekiz yıllık tecrübe*).

Yeterlik alanları oldukça kapsamlı ama herkes farklı bir şey anlarsa ne olacak? doğru değerlendirme yapamaz ki. Burada daha bu alanları birbirine karıştıran çok öğretmen var. Açıklayıcı ve yazılı bir rehber lazım. Yeterliklerin ölçütleri de açıklanmalı yani göstergeler net olmalı. (P3 *onaltı yıllık tecrübe*).

Şimdi sen söylüyorsun ya ben şu neydi bu neydi diyorum bazen içimden. Mezun olalı yıllar oldu. Öğretmen kılavuz kitabı olsa, içinde amacı, süreci, sonucu anlatsa... Öğretim programlarının eski halinde vardı buna benzer bir şey ve bana çok yardımcı oluyordu. (T3 *ondört yıllık tecrübe*).

Evet hatta bu rehber eğitimlerde bile kullanılabilir. Herkesin elinin altında olur. Değerlendiren de değerlendirilen kişi de nasıl bir model olduğunu bilir ve daha istekli uygular. Bu rehber ekleme istediğimiz ya da çıkartmak istediğimiz şeyler olabilir bunlar bile bize sorulabilir. Rehberin anlaşılır ve kullanışlı olması çok önemli. (T21 *üç yıllık tecrübe*).

Since the participants talked about the need for a guide, the researcher asked how the content of this guide should be. In a group where focus group interviews were held, participants stated that the importance and necessity of teacher evaluation should be

explained in the guide of this model. The participants (T9, T13, T21, P1, E3) explained that this explanation would increase the willingness to use the model, increase the belief in the model and apply the model in a more motivating way. When this suggestion was presented to other groups, the participants found it appropriate to include this section in the guide for similar reasons. A teacher (T13 twelve years of experience) made the following statement on this subject "*Teacher evaluation is a difficult concept to accept. In order not to be afraid of evaluation, you need to know what it means and its importance. This importance should definitely be included in the guide. (Öğretmen değerlendirme kabul etmesi zor bir kavram. Değerlendirmeden korkmamak için onun ne demek olduğunu ve önemini bilmek lazım. Bu önem rehberde mutlaka yer almalı.)*"

Participants emphasized that the purpose of this model should be included in the guide. In all focus groups, it was mentioned that including the purpose of the evaluation as for professional development would give more confidence to the reader of the guide and that knowing the purpose is also vital in order to apply the model correctly. One teacher expressed this importance as follows.

The fact that the purpose is in a written form makes it clear to everyone. This evaluation is not made to punish. They can easily apply it. Also, I should know my purpose so that I can understand why these methods are carried out the way they are. (T23 four years of experience)

"Amacın yazılı bir biçimde yer alması herkes için bir açıklık getiri. Bu değerlendirme ceza vermek için yapılmıyor. Rahatlıkla uygulayabilirler. Bir de amacımı bileyim ki bu yöntemler süreçler neden böyle yürütülüyor anlayabileyim." (T23 dört yıllık tecrübe)

In one focus group, the participants explained that it is important to include the development process of this model in the guide. The participants (T9, T13, T21, T22, T23, P1, E3) emphasized that knowing that they, that was, the people who would be affected by the evaluation, developed this model was an indication that the model meets the needs of the institution and will increase the willingness to implement this model. When this suggestion was presented to other groups, the participants found it appropriate to include this section in the guide for similar reasons. A teacher explained the importance of including the development process of the model in the guide as follows.

I've been in this process from the beginning, and it's great that this model reflects our views. It meets all our needs and is unique to this institution. Of course, the things that don't work are updated over time, but every teacher in this school applies it with peace of mind. This should be added to the guide so that everyone knows. Even a teacher who has just started the institution should know. (T10 four years of experience)

Başından beri bu süreçteyim ve bu modelin görüşlerimizi yansıtması harika bir şey. Bizim tüm ihtiyaçlarımızı karşılıyor ve bu kuruma özgü. Tabi ki işlemeyen yerler güncellenir zamanla ama bu haliyle bu okuldaki her öğretmen gönül rahatlığıyla uygular. Bunun rehberde eklenmesi lazım ki herkes bilsin. Hatta kuruma yeni başlayan bir öğretmen de bilsin. (T10 dört yıllık tecrübe).

In all focus groups, participants indicated the importance of including all qualification areas with subdimensions and even criteria in the guide. These explanations were needed especially because of the importance of developing clarity regarding the qualification area to be evaluated and the importance of everyone having the same opinion about this area of qualification. The participants (T5, T11, T15, T23, T25, T26, T29, T32, P1, P5, E3, E6) also stated that behavioral indicators, that was, criteria or standards explaining these qualification areas, would be useful for everyone to understand in the same way and to develop a common understanding of evaluation. An expert has expressed this importance.

For the effectiveness of the evaluation, it is very important to use a common language, that is, to evaluate with the same terminology. This can only be achieved by defining behaviors. What are the teacher behaviors that explain the qualification? If we know this, we will make the right assessment. This must be added to the directory. (E3 five years of experience)

Değerlendirmenin etkililiği için ortak dil kullanımı yani aynı terminoloji ile değerlendirme yapmak çok önemli. Bu da ancak davranışları tanımlayarak olur. Yeterlikleri açıklayan öğretmen davranışları nelerdir? bunu bilirsek doğru değerlendirme yaparız. Mutlaka rehberde eklenmelidir (E3 beş yıllık tecrübe)

The participants (T3, T5, T9, T15, T23, T25, T26, T29, T32, P1, P2, P5, E3, E6, E7) also stated that the inclusion of brief information on how to use each data collection tool developed would raise awareness about these tools. Some of the participants (T3, T9, T15, T23, T26, T32, P1, P5, E3, E6) stated that the training in which the use of these tools would take place could also be briefly mentioned. One teacher explained

(T15 *four years of experience*), “All the tools should be in the guide. They should be introduced briefly, and it can even be mentioned in the guide that the use of these tools will be provided through trainings (Değerlendirme araçlarının hepsi rehberde olsun. Kısa kısa tanıtılsın, hatta bu araçların kullanımının eğitimlerle sağlanacağından bile bahsedilebilir rehberde.)”

Participants stated that it was important to mention that the process was carried out consistently and objectively. Furthermore, most of them (T8, T9, T11, T22, T23, T24, T26, T28, T29, T32, P1, P3, P5, E1, E3, E7) stated that it was important to mention that the results will be given with the principles of openness and confidentiality. One teacher (T26 *five years of experience*) expressed this requirement as follows: “If the first problem that comes to our mind here is objectivity and confidentiality, this may come to everyone’s mind. We should clearly state how we will achieve this in the guide. (Burada bizim aklımıza gelen ilk problem, objektiflik ve gizlilik ise bu herkesin aklına gelebilir. Bunu nasıl sağlayacağımızı açıkça belirtmeliyiz rehberde.)” Another teacher stated:

Everything should be clear, how to give feedback, how to keep this information confidential, and how to be objective. What if it’s not consistent? Will I be given the right to express myself? It is so important that all of them should be included in the guide one by one. (T8 *eighteen years of experience*)

Her şey net olmalı geri dönüt nasıl verilecek, bu bilgiler nasıl gizli tutulacak, objektif nasıl olunacak. Tutarlı olunmazsa ne olacak? Kendimi ifade etme hakkı verilecek mi? O kadar önemli ki hepsi tek tek yer almalı rehberde. (T8 *on sekiz yıllık tecrübe*)

While examining the schedule table in general, the participants mentioned that each data collection tool mentioned in the table was essential. They stated that these tools should be developed in line with detailed criteria, and these tools should be prepared and presented to them. The participants stated that they were willing to give their contribution to the model itself in the examination of the tools. Many participants also added that these tools would be more functional if they were piloted. A teacher explained this wish as follows.

All these tools are very important. The model must include the final version of these forms. While developing these, they need to be designed with very detailed criteria. If forms are prepared with these criteria, and they are sent to us for ideas, I can express my views. (T3 fourteen years of experience)

Bu kullanılacak araçların hepsi çok önemli. Modelde mutlaka bu formların son hali olmalı. Bunların geliştirilirken çok detaylı ölçütlerle tasarlanması lazım. Bu ölçütlerle formlar hazırlanırsa ve bizlere fikir alma amaçlı yollanırsa görüşlerimi söyleyebilirim. (T3 on dört yıllık tecrübe).

Another teacher stated:

It is important to use these tools, that is, to try. Forms get better as they are used, and the developed forms should be included in the model. Maybe we can't evaluate some criteria. Then that criterion should be changed, for example. A pilot may be necessary. (T24 five years of experience)

Bu araçları kullanmak yani denemekte önemli. Formlar kullanıldıkça gelişir ve geliştirilen formlar model de yer almalı. Belki bazı ölçütleri değerlendiremeyiz. O zaman o ölçüt değiştirilmeli mesela. Bir pilot uygulama gerekli olabilir. (T24 beş yıllık tecrübe)

4.3 Third Part of The Study

In the third part of the study, the guide of the model and the data collection tools to be used in the model were developed in line with the decisions made after the focus group discussions. The guide and all data collection tools were shared with all classroom teachers, heads of departments, math and science teachers, principals, and experts working at the school via e-mail. In addition, in this section, the views of the participants about the developed guide and the changes made, as well as how the validity of the data collection tools was ensured, were explained.

4.3.1 The Guide and The Tools of The Teacher Evaluation Model

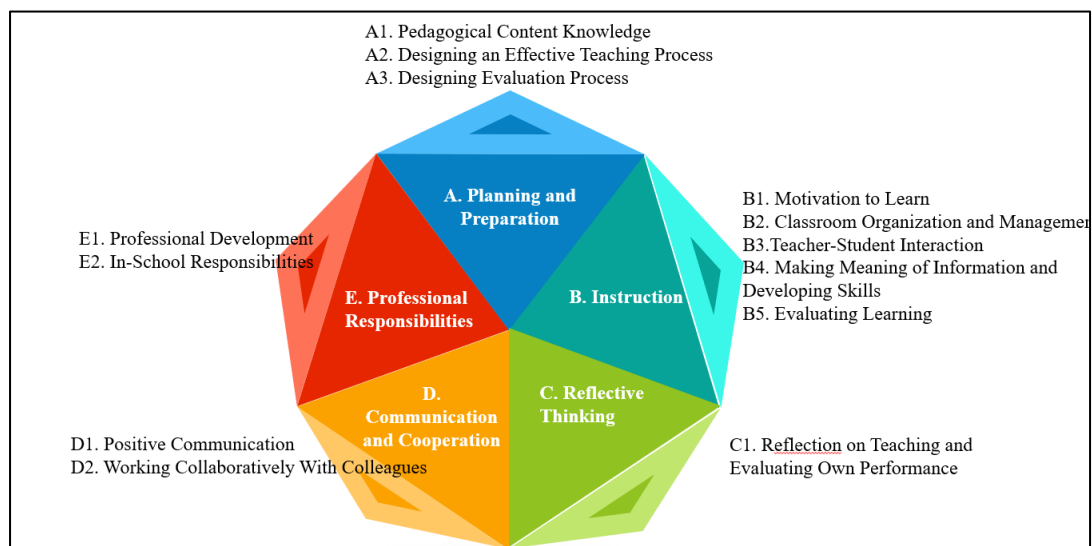
Participants emphasized that there should be a section describing the importance, purpose, and development process of the model in this guide. Based on this suggestion, an introductory section was created in the guide, explaining the importance of teacher evaluation and the rationale and the process of designing a participatory evaluation specific to a school. While creating this introductory part, the opinions of the

participants were taken as a basis, and many sources in the literature (Danielson & McGreal, 2000; Kane et al., 2014; Marzano & Toth, 2013; OECD, 2009b; OECD, 2013a; Ofsted, 2018; Stronge, 2018; TEDMEM, 2018) were also examined.

Participants emphasized that a teacher evaluation model should be both specific and comprehensive in terms of teacher qualifications and subdimensions of the qualifications. The importance of a criteria-based teacher evaluation model that covers and reflects effective teaching behaviors and includes these criteria in data collection tools in a comprehensive manner are the issues emphasized by the participants in both data collection processes. For this purpose, each qualification area, corresponding subdimensions, and criteria were developed by reviewing well-cited teacher evaluation models in the field (Clayton, 2017; Danielson, 2013; Marzano & Toth, 2013; Marzano & Simms, 2014; Shinkfield & Stufflebeam, 1997; Sloat et al., 2017; Toch & Rothman, 2008), books (Bender, 2005; Danielson, 2007; Danielson & McGreal, 2000; Deiro, 2005; Egan, 2010; Geng et al., 2019; Graham et al., 2015; Graham & Berman, 2018; Hattie, 2012; Hernandez & Endo, 2017; Jones et al., 2006; Kennedy, 2005; Kohn, 2000; Leighton, 2020; Marshall, 2013; Robinson & Aronica, 2016; Shulman, 1986; Stronge, 2018; Tucker & Stronge, 2006; Walsh & Sattes, 2015; White, 2016) and articles (Black & William, 1998; Brookhart, 2020; Kennedy, 2016; Liberman, 2000; TEDMEM, 2018) written by the proponents of teacher evaluation. With the help of the researched sources, qualification areas and subdimensions that the participants clarified with the help of the focus group interview and the criteria for each sub-qualification are given in detail in the model. Teacher qualification areas and sub-areas included in the guide of the model are presented in Figure 4.2.

Figure 4.2

Teacher Qualification Areas and Subdimensions Included in The Guide of the Model



In the focus group interviews, the participants stated that the importance of the training for the people who would make the evaluation and be evaluated, the content of this training, and how this training would be done should be explained in the model. Both in one-on-one interviews and in focus group interviews, it was stated that training should be given to the people who will evaluate and be evaluated at regular intervals about the purpose of the model, its process, the effective use of data collection tools, and the reporting process. In the focus group interviews, it was also emphasized that this training, especially the training to be given for classroom observation, should be practical. In these interviews, the participants suggested that the scenario examples describing the classroom environment could be examined in the training. These recommendations are added to the descriptions of the model.

On the other hand, the schedule, which includes the date of the assessments, who will conduct the assessments for which competency areas, and data collection tools to be used in these assessments, was updated in line with the suggestions from the focus group interviews. Detailed information on the use of data collection tools in this schedule was added to the model in line with the opinions of the participants.

In addition, consistency, and objectivity, which were important issues frequently mentioned in both one-on-one interviews and focus group interviews, were also included in the model. The need for detailed explanations of these issues emerged in the focus group discussion. In line with the suggestions of the participants, things to be done to provide consistency and objectivity were added to the model. Finally, detailed information about the openness and confidentiality of the evaluation results, which was especially emphasized by the participants in the focus group interviews, is included in the explanations section.

After the participants approved the data collection tools in the schedule table of the initial model, four different data collection tools were prepared by the researcher. These tools are, respectively, the lesson plan evaluation form, the classroom observation form, which is used by conducting interviews before and after (Pre-Observation Interview Form, Observation Form, and Post-Observation Interview Form), the communication and cooperation evaluation form, and the professional responsibilities evaluation form.

Lesson Plan Evaluation Form

While preparing the lesson plan evaluation form, the qualification area descriptions and behavioral criteria defined in the model were taken as a basis. It was aimed to determine the adequacy of the lesson plan related to this evaluation form in terms of the specified criteria. Within the scope of this model, the lesson plans are planned to be evaluated by the school principal, vice principals, classroom teachers, head of departments, and teachers from different subject matters (science and mathematics teachers). According to the difference of the evaluators, the criteria in the forms were differentiated when necessary. For example, since the principals did not attend the classes for a long time and not all of them were classroom teachers, the form prepared for them did not include criteria such as “include preliminary information, scientific accuracy of the information, elimination of misconceptions.” In this form, the evaluators would be asked to assess each statement on a four-point rating scale: very well developed, basic level, sufficient, exemplary. They would be asked to mark the

“Not included” (fI) category if the given situation was not related to the relevant course content or if it was not possible to observe it in the relevant course content. This form also includes the meanings of the categories corresponding to the effectiveness level of the lesson plan.

Pre-Observation Interview Form, Observation Form, and Post-Observation Interview Form

Evaluation of the teaching process consisted of three parts. In the first part, there was an interview form to be used to collect information about the class and the instruction process with the teacher before observing the classroom. In the second part, there was an observation form to be used in order to observe the instruction and to determine its sufficiency in terms of the specified criteria. In the third part, there was an interview form that will be used to share the observation results with the teacher of the lesson after observing the lesson process, to get the teacher’s views on the process, and to evaluate the teacher’s reflective thinking.

"Teaching Process Pre-Observation Interview Form" was to be filled in before the lesson observation by meeting with the teacher of the lesson to be observed. With the help of this interview form, information about the class to be observed (the general profile of the class, information about the students who stand out with their positive and negative characteristics in the class, and the precautions to be taken) and information about the learning and teaching process (the purpose of the lesson plan to be applied, the purpose of the lesson plan to be applied in the lesson to be observed) section(s) were designed to be collected. Within the scope of this model, this form, which is used before the lesson process, was planned to be used by the school principal, administrators, class teachers, head of departments, and teachers from different subject matters (science and mathematics teacher).

On the other hand, the observation form was developed to observe the instruction process based on the qualification area descriptions and behavioral criteria defined in the model. Through the use of this evaluation form, it was aimed to determine the effectiveness of the course process in terms of the specified criteria. Accordingly, the

participants would be asked to assess each criterion on a four-point scale with labels: highly developed, basic level, sufficient, exemplary. In addition, the "Not Included" option was included for the situations not related to the relevant course process, not possible to observe during the course, or not remembering what was done during the course. Directions included advising participants to read the form carefully before the observation to facilitate their decision-making regarding the relevant criteria during the lesson observation and note-taking process about the situations they observed during the lesson observation. Within the scope of this model, the instruction process was planned to be evaluated by the teacher themselves, the school principal, the vice principals, the classroom teachers, the heads of the departments, and the teacher from different subject matters (science and mathematics teachers).

After observing the instruction process, the teaching process of the teacher was planned to be evaluated by the teacher themselves. For this purpose, topics such as the effectiveness of the course, the use of methods and techniques, and the attractiveness of the course were evaluated together with the teacher. In the continuation of this section, a short form used to evaluate the teacher's reflective thinking was also included. Teachers would be asked to self-assess on a four-point rating scale: "Never," "Partly," "Mostly," and "Completely." Within the scope of this model, the instruction process was designed to be evaluated by the school principal, the vice principals, the classroom teachers, the heads of the departments, and the teacher from different subject matters (science and mathematics teachers).

Communication and Collaboration Evaluation Form

While preparing the communication and collaboration evaluation form, the qualification area descriptions and behavioral criteria defined in the model were taken as a basis. Through the use of this evaluation form, it was aimed to evaluate the situation of the teacher in terms of communication and collaboration in line with the criteria specified. Accordingly, the evaluator would be asked to rate the frequency of the situation on a four-point rating scale: "Never," "Rarely," "Often," and

"Completely." Within the scope of this model, this form was planned to be used by the classroom teachers and heads of departments.

The Professional Responsibilities Evaluation Form

While preparing the professional responsibilities evaluation form, the qualification area descriptions and behavioral criteria defined in the model were taken as a basis. Through the use of this evaluation form, it was aimed to evaluate the professional responsibility of the teacher in line with the criteria specified. Accordingly, the evaluator would be asked to rate the frequency of the situation on a four-point rating scale: "Never," "Rarely," "Often," and "Completely." Within the scope of this model, this form was planned to be evaluated by the school principal, vice principals, heads of departments.

4.3.2 Opinions of the Participants Regarding the Guide of The Model (Research Question 6)

The final version of the prepared guide was shared with 60 classroom teachers, 11 principals and 7 experts working at the school via e-mail. First of all, a short information about this research was given and the development process of the model was explained. The participants were asked to examine the guide of the model in detail, and evaluate in terms of scope, comprehensibility and feasibility. They were also asked to add whether they support professional development and, if any, what can be suggested to support professional development. They were asked to write their opinions as comments in the relevant sections of the guide. It was also requested from these people to clearly state the places that need to be changed, together with their suggestions.

32 of the teachers, 5 of the administrators and 6 of the experts returned the mail. When the opinions about the model are examined, the participants stated that the model fully supported professional development. They also mentioned that the guide of the model was clear, understandable, comprehensive, applicable and qualification criteria include all kinds of behaviors that a teacher can display while performing their

profession. In addition, the participants also stated that they were willing to implement this model because it reflects their own views and is designed to meet the need. 3 teachers and one expert suggested to use more understandable and clear words instead of some words. For example, two teachers (T23, T30) suggested using the phrase "teacher shares" instead of "teacher tells", while another teacher (T29) and an expert (C2) suggested using the word "daily life" instead of "real life". In addition, the participants expressed their opinions about grammar and punctuation.

It is thought that the statements of these participants, especially teachers, that they are willing to be evaluated within the scope of this model, are important. Some of the participants expressed their views as follows.

"I have explained my views and suggestions since the beginning of the model. I was very happy to see that all of them were reflected. It is very comprehensive and understandable. I wish the principals to take it into consideration and implement it in the school as soon as possible...(Modelin başından beri görüş ve önerilerimi açıklamıştım. Hepsinin yansıtıldığını görmekten çok mutlu oldum. Gayet kapsamlı ve anlaşılır. Yöneticilerin dikkate alıp okulda bir an evvel uygulanması dileğiyle...) teachers "(T3 *fourteen years of experience*)

"I just made some spelling corrections in a few places. Apart from that, the guide is great. It is possible to provide professional development with this model. So much effort. Let these efforts not be wasted and implemented as soon as possible (Bir kaç yerde imla düzeltmem oldu sadece. Onun dışında rehber harika olmuş.Bu modelle mesleki gelişim sağlamak mümkün. Ne çok emek var. bu emekler boşa gitmesin ve bir an evvel uygulansın)" (T7 *fifteen years of experience*)

"I've been reading, but I couldn't find a place to fix it. I think of something to add, then I look at the following chapters, the same thing. I like it very much (Okuyorum ama düzeltecek yer bulamadım. Aklıma bir şey geliyor ekleyeyim diyoru sonra bakıyorum devam eden bölümlerde var aynı. Çok beğendim)" (E2 *altı yıllık tecrübe*).

4.3.3 Validity Evidence of Teacher Evaluation Tools (Research Question 7)

For the validation of the tools, feedback was gathered through emails, and then pre-piloting process was conducted. First of all, a short information about this research was given and the development process of the tools was explained. The final version

of the guide of the model was also sent in the e-mail and it was stated that the guide could be examined for the parts that were not understood while expressing opinions about the tools. In the email, the researcher asked participants to evaluate the data collection tools in the model in terms of qualification areas, measurability of the criteria, and applicability of the tools. The tools were shared with 30 classroom teachers, 11 principals, 7 experts, 5 math and 5 science teachers working at the school via e-mail. 10 teachers, 1 mathematics, 2 science, 3 principals, and 5 experts have expressed their opinion via e-mail that the tools are comprehensive, that the behaviors can be measured and that they are applicable. 7 classroom teachers, 1 mathematics teacher, 1 science teacher, 4 experts and 4 principals gave their detailed opinions on the items and categories in the tools. These views and the arrangements are detailed in the continuation of this section.

For the pre-pilot of the data collection tools, tools were shared with 10 classroom teachers, 6 administrators, 5 mathematics and 5 science teachers working at the school via e-mail. Some of the participants have agreed to participate in the pre-pilot. Therefore, tools were pre-piloted with 5 classroom teachers, 1 math teacher, 1 science teacher, and 2 principals. During the pre-pilot process, the participants were asked to use the tools and to express their opinions on the clarity and evaluability of the expressions in these tools. Due to the covid 19 epidemic, the practices for the course observation process were stopped at the school where the research was carried out. Therefore, in this process, the following tools were not piloted: “Pre-Observation Interview Form,” “Observation Form,” and “Post-Observation Interview Form.” Only the teachers self-evaluated their instructional processes with the help of the observation form. The results of the pre-pilot process and the corrections made are given in detail in the next sections. Table 4.25 lists the teacher evaluation tools designed in this study, participant groups that tools were sent for feedback and/or pre-pilot, and the process.

Table 4.25*Tools, Evaluators, and Validation Process*

<i>Tools</i>	<i>Participant groups</i>	<i>Validity process</i>
Lesson Plan Evaluation Form	Head of the department Classroom teachers Math, Science teacher Principals	Feedback from the participants Pre-pilot
Pre-Observation Interview Form Observation Form Post-Observation Interview Form	Head of the department Classroom teachers Math, Science teacher Principals Teacher herself	Feedback from the participants Pre-pilot (only teacher herself)
Communication and Cooperation Evaluation Form	Head of the department Classroom teachers	Feedback from the participants Pre-pilot
Professional Responsibility Evaluation Form	Head of the department Principals	Feedback from the participants Pre-pilot

4.3.3.1 Feedback About the General Structure of The Tools*Regarding the items on the scale*

Three experts (E2, E4, E6) and one teacher (T11) stated that the items containing more than one judgment and connected with the conjunction "and" should be reviewed because it would not be understood for which of these judgments evaluators made the evaluation. For this purpose, such items have been divided and added to the scales as two separate items. For example, the "identification and elimination of misconceptions" items are divided into two different items as "identification of misconceptions" and "elimination of misconceptions." Two experts (E3, E6) and a principal (P9) stated difficulties matching some criteria with the response categories. As a result, some of the items and response categories changed, making items more easily responsive. For example, while the expression "somewhat" in the communication and cooperation evaluation form was replaced with the expression "rarely," the expression "quite a lot" was replaced with the expression "often" so that behaviors can be evaluated under these categories.

4.3.3.2 Lesson Plan Evaluation Form

Feedback regarding category definitions

A teacher (T5) stated that some sentences in the description of the "highly developed" category were not intelligible. When the explanation under the "highly developed" heading was examined, it was realized that if the sentences were corrected grammatically, the explanation would be more meaningful. The necessary grammatical corrections were made. An expert (E2) suggested that including a statement like "a plan was made as expected to be" would distinguish this category from others clearly. When this category was examined, it was determined that such a statement was needed. So that the statement, "The situation can be realized ideally within the framework of the planning and explanation," was added. One of the teachers (T11) and two experts (E1, E4) agreed that the "adequate" category in the lesson plan evaluation form was not fully understood and should be written more clearly. The researcher simplified the statements in this category.

Feedback regarding the items on the scale

An expert (E4) and a principal (P3) stated that some assistant principals would have difficulty in evaluating "the suitability of methods and techniques" and items related to learning styles since they did not participate in a course for a long time. They suggested that these items could be removed. A science teacher (T33) and math teacher (T34) stated that they might not know whether the material or resources to be used in the course process are exciting or intriguing because they do not teach in this age group. They stated that it was also difficult to say anything about the appropriateness of methods and techniques for this age group because they teach older children. In addition, two classroom teachers (T3, T11) stated that it would not be correct for subject matter (math and science) teachers to express their opinions on these items. Therefore, the specified items have been removed from the forms of the evaluators concerned. An expert (E2) stated that "the suitability of the materials and resources to be used in the course process for the subject and outcome" should also be evaluated. Another expert (E6) stated that an item could be added to this section

regarding whether the planned process was realistic or not. When the experts' opinions were examined, these two items were considered essential and added to the forms.

Pre-Pilot Results

After the necessary arrangements were made in this form, the final version of the form was sent to the participants via e-mail. Two heads of department (T11, T5), two classroom teachers (T14, T18), a science teacher (T35), a mathematics teacher (T36), and two assistant principals (P5, P6) were asked to evaluate a lesson plan of their choice using this form. One of the participants (T11) and a teacher (T18) stated that it was important to include the outcome along with the subject in the first section of this form. When the reason for this change was asked, the teacher (T18) stated that when she evaluated other lesson plans prepared for the same outcome, an effective plan could be shown as an example and used by other teachers. The head of the department (T11) explained that it was essential to search by learning outcome if these forms were uploaded to the school portal. In line with these explanations, the title of "outcome" was added next to the title of the subject. A mathematics teacher (T36), science teacher (T36), and a principal (P5) stated that it was also essential to use the time effectively to carry out the planned activities. Considering this suggestion, the item "containing a realistic time frame that will allow the outcome to be realized" was added. The other participants who carried out the piloting process did not make any suggestions about the form and stated that the form could be used as it is.

4.3.3.3 Pre-Observation Form, Observation Form, and Post-Observation Form

Feedback regarding category definitions

An expert (E4) stated that using the expression "mostly" instead of the "limited" term mentioned in the "adequate" category in the teaching process evaluation form would be a better label.

Feedback regarding the items on the scale

One expert (E2) stated that it would be appropriate to remove the item "Taking care of the students in need" from the "Affective Support" heading. A principal (P8) and a

teacher (T23) stated that it would be more appropriate to include the item "Respect and understanding of their behaviors towards students" under the title of " Effective Communication" in the dimension of teacher-student interaction instead of " Affective Support" title. The specified corrections have been made.

One of the teachers (T3) stated that the deepening process mentioned in the item "Deepness of explanations about the knowledge (subject/concept/relationships, etc.)" under the heading "Student's Interaction with New Knowledge" should be in line with the limitations of the program. For this purpose, the expression "in line with the program limitations" has been added to this statement. Two experts (E2, E6) stated that the appropriateness of the questions in the expression "Questions about new learning during the course" under the heading of "evaluation" was also important, this statement could be stated more clearly, and there is no need for the expression "during the course." Therefore, this expression has been changed to "Asking appropriate questions for new learning."

Feedback regarding the post-observation interview form

A teacher (T22) stated that questions like "Was the lesson plan effective in the post-observation interview form? Did you implement the plan effectively?" can also be asked after asking the question "How would you evaluate your course in general?". A teacher (T29) and a principal (P3) stated that the expression "Are there any situations for this course that you think will facilitate learning and make the knowledge/skills permanent for a long time?" in the post-observation interview form was not fully understood. They also noted that this expression could be asked in a more precise and sincere language. For this purpose, the statement changed to "Can you tell me what you did for this course to facilitate the relevant learning or ensure that the knowledge/skills are long-lasting?". An expert (E4) mentioned that "Which part was the most attended and the most interesting part of the students?" should be added to "Do you think the lesson was interesting for the students? How was the participation of the students in the lesson?" questions, to gather more detailed answers.

Pre-Pilot Results

“Pre-Observation Interview Form, Observation Form, and Post-Observation Interview Form” could not be piloted by the participants due to Covid-19 restrictions. Only the teachers self-evaluated their instructional processes with the help of the observation form. Two classroom teachers (T14, T18) tested the form by evaluating themselves after the classroom sessions. The teacher (T14) noticed that some of the statements in the form did not contain verbs and added action sentences at the end of the two items to make them similar to the other items. The specified change was reflected in the form. The other teacher (T18) stated that the two items in the form could cover each other. When the necessary explanation was made, it was confirmed by the teacher that these two items were different from each other. Both teachers stated that they could not observe all the items in a single lesson, but they could evaluate most items when they evaluated a few different courses.

4.3.3.4 Communication and Cooperation Evaluation Form

Feedback

One of the experts (E1) criticized the categories and stated that it would be more appropriate for these behaviors to change the categories "Never," "Slightly," "Quite," and "Verily" to "Never," "Rarely," "Often," "Almost always." These corrections were considered so that the behaviors could be evaluated easily. An expert (E6) stated that the expressions under the heading "Working in Collaboration" also included working in harmony, so the phrase "Working in harmony with his colleagues." can be omitted. One teacher (T11) mentioned, "Working in harmony with his colleagues." statement remained very general compared to other statements and could be deduced. By making these two corrections, the "Working in Collaboration" area was reorganized and "Working in harmony" item was removed.

Pre-Pilot Results

After the necessary arrangements were made in this form, the final version of the form was sent to the participants via e-mail. Two heads of department (T11, T5), two

classroom teachers (T14, T18) piloted the forms. The teacher (T14) suggested that it can be written more precisely by explaining the support departments in the statement using parentheses. Therefore, the names of the relevant departments have been added to this item in parentheses. The teacher (T11) noticed "Colleagues help out when they need it." statement was not directed at the evaluator. So that this statement changed to "She helps her colleagues when they need it." The other participants who carried out the piloting process did not make any suggestions about the form and stated that it could be used as it is.

4.3.3.5. Professional Responsibility Evaluation Form

Feedback

Two principals (P3, P7) stated that they would have difficulty expressing their opinions about "Following students' progress by collecting evidence for their academic development" and "Following students' progress by collecting evidence for social development" statements. Therefore, these two items could be evaluated more effectively by the heads of the groups; for this purpose, the form states that these items will only be evaluated by the head of the group and the school principal.

Pre-Pilot Results

After the necessary arrangements were made in this form, the final version of the form was sent to the participants via e-mail. Two heads of department (T3, T5) and two principals (P5, P6) piloted the forms. One of the principals (P6) stated that the statement "She is improving herself in line with the feedback in the development plan" was a general expression, and it would have taken a long time to happen. So that this statement was changed to "The teacher organizes his/her work in line with the feedback in the development plan." The other participants who carried out the piloting process did not make any suggestions about the form and stated that the form could be used as it is.

4.4 Summary of Findings

Current Practices and Problems about Teacher Evaluation

In the school where the study was conducted, the purpose of teacher evaluation was not clear, and evaluations were not made with certain standards or criteria. There was no certain evaluation schedule, tools were not validated, and the data sources (average achievement scores, lesson plans, and informal principal observation) were limited. All these problems indicated that the teacher evaluations conducted at this school are not systematic. Due to these unsystematic evaluations, the participants did not have sufficient information about teacher evaluations. The uncertainty they experienced prevented them from trusting the evaluations. It is one of the important findings of the present study that the teachers were given only summative feedback in general meetings. It was also revealed that teachers did not receive any formative, constructive, or individual feedback about the effectiveness of their teaching.

The absence of clear criteria and a functional reporting process for sharing evaluation results and feedback in the school, where the research was conducted, adversely affected both the usefulness of the evaluations and the defensibility of the results for the evaluators. In addition, the existence of an environment that supports competition among teachers in the school also negatively affected accuracy. Teachers are compared with the average scores of students in school-wide exams. The teacher who teaches students with a high average score is appreciated more than other teachers. In other words, judgments are made about the effectiveness of the teacher according to the average success of the students in the school-wide exams. When this situation is examined, even if it is not written or explicitly stated it can be said that teacher evaluation is made to increase student success. This process of evaluation was not reflecting the complexity of teaching and how the teacher supports the development of students in other areas such as the affective domain. Decisions made in favor of experienced teachers and individual relations between teachers and teachers or between teachers and principals also affected accuracy as they cause biased

evaluations. This privileged situation, which is given to experienced teachers, causes teachers to develop a belief that there is an unfair evaluation.

Since this research aimed to develop a teacher evaluation model aiming professional development, the opinions of the participants were also taken about the professional development studies carried out at the school. It has been revealed that in-service training seminars in the school are carried out based on the topics chosen by most of the teachers or on popular topics. This training is usually given to all teachers collectively and the teacher rarely participates actively. General training given to the large masses is interpreted as a problem being ineffective since they do not meet individual needs, what is learned is not permanent, and the participants cannot participate.

Teacher Evaluation Model Needed

Participants stated that the purpose of the model should be the evaluation and development of teacher qualifications (communication and cooperation, teaching, responsibilities towards school, lesson preparation, monitoring and evaluation, professional development). The teachers mentioned the importance of being fair, having positive communication skills, having teaching experience, empathizing, competent in teacher, knowledgeable about the school, and communicating effectively, whom the principals, classroom teachers, teachers from different branches and the teacher themselves could evaluate. Participants also stated that evaluators should receive training on how to use data collection tools and what to pay attention to in the evaluation process before applying the model.

The participants stated that the problems related to the observations were the lack of systematic observations at school, the teachers' lack of knowledge about the observation results, and the lack of evaluation criteria. For this reason, the participants stated that there was a need for evaluations made with valid observation forms prepared in line with certain criteria as a data collection tool in the evaluation. In addition, it was explained that there was a need for interviews before and after the observation.

The participants stated that the evaluation results should be shared only with the teacher in accordance with the confidentiality principle. Participants also explained that the evaluation results should include both strengths and weaknesses, that a report containing feedback and a teacher development plan based on multiple measures would be functional, and that the report should be provided at the end of each training period to give teachers enough time to improve.

It was also emphasized by the participants that in order for the evaluation results to be unbiased, the data should not be obtained objectively, that was, decisions should not be made based on personal relationships. Teachers working in this institution stated that the behaviors of experienced teachers were tolerated depending on their good relations with the principals. On the other hand, the behavior of a novice teacher was constantly examined and tried to be found incomplete. The participants stated that it is important to evaluate all teachers by following the same process, regardless of their professional experience.

It has been revealed that in-service training seminars at school are carried out on topics chosen by the majority of teachers or on popular topics. This training is usually given to all teachers collectively and the teacher is rarely actively involved. General education given to large masses is interpreted as a problem of being ineffective because it does not meet individual needs, what is learned is not permanent and participants cannot participate. Participants mainly suggested that teacher evaluation results should be used when planning professional development activities.

Model Development

In the second part of the study, the researcher developed an initial model in line with the opinions obtained from the first part of the study. The draft model was introduced to the participants and their opinions were taken through focus group discussions. As a result of the focus group discussions, the qualification areas and sub-dimensions planned to be included in the model were changed and arranged. While the participants stated that each qualification area has equal importance and there is no need to prioritize, it was also emphasized that no difference should be made for novice and

experienced teachers. It was also decided in these meetings that the trainings to be given to the evaluators should be applied and regularly repeated.

When the previous interviews were examined, it was determined that the evaluation of the lesson plans should not be done because it would not fully reflect the instruction process. In the focus group interviews, the participants were asked what they thought about the use of the lesson plan in such a model. Participants stated that lesson plans can be used as data source as long as they are not seen as a single resource and the stages specified in this model are carried out.

Participants were asked to evaluate the feasibility of this model. Participants stated that the specified times were appropriate for the implementation of this model. In the focus group interviews, most of the participants stated that their own views are included in this model and that it is a collaborative model, and they explained that this situation makes them feel valuable. Participants stated that they are willing to be evaluated within the scope of this model, in which their views are reflected, and their needs are considered.

When the participants were asked questions about the different dimensions of the initial model and especially when asked to evaluate its feasibility, they stated that there was a need for a guide in which this model was explained in more detail. The common view of the participants in all groups is that a guidebook can be prepared in which the purpose of this model, areas of qualifications, the tools to be used, the timeline, how the process will be carried out and how the teachers will be informed, can be prepared and this guide is important in terms of ensuring unity in the application of this model. For this purpose, it was decided to prepare a guide at the end of the meeting. Participants also stated that data collection tools should be prepared and presented to them in line with detailed criteria, and they were willing to contribute to the development of these tools and the model itself. Many participants also added that these tools would be more functional if piloted.

In the third part of the study, the guide of the model and the data collection tools to be used in the model were developed in line with the decisions taken after the focus group

discussions. All qualification areas, sub-dimensions and criteria were given in detail in the guide. The purpose of the model, the importance of teacher evaluation and the development stages of this model are also included in the guide. It is also explained in detail how the evaluations will be made with the help of the model and how to ensure objectivity, consistency and confidentiality. Data collection tools were developed with the help of the criteria in the guide. The guide and all data collection tools were shared with all classroom teachers, department heads, mathematics and science teachers, principals and experts working at the school via e-mail. In line with the opinions received, additions were made to some parts of the guide, words were arranged, and some expressions were made more understandable. In addition, pre-pilot applications were carried out in this section to ensure the validity of the data collection tools.

CHAPTER 5

DISCUSSION AND IMPLICATIONS

In this chapter findings are discussed under the following headings: current practices and problems of teacher evaluation, teacher evaluation model needed, and development of a school-based, participatory teacher evaluation model for professional development. Implications for educational practice and further research are also explained in the chapter.

5.1 Current Practices and Problems About Teacher Evaluation

In the school where the study was conducted, the purpose of teacher evaluation was not clear, and evaluations were not made with certain standards or criteria. There was no certain evaluation schedule, tools were not validated, and the data sources (average achievement scores, lesson plans, and informal principal observation) were limited. All these problems indicated that the teacher evaluations conducted at this school are not systematic. The present study has shown that those working in this school need a systematic teacher evaluation that has a clear purpose, includes explanatory criteria, has a certain evaluation schedule, and uses valid data collection tools. All these problems have shown that the teacher evaluations in this school are far from being systematic.

Due to these unsystematic evaluations, the participants did not have sufficient information about teacher evaluations. The uncertainty they experienced prevented them from trusting the evaluations. Policy statements and/or guidelines that explain the purpose of the evaluation system, how the system will be used, and the data collection and reporting process ensure that evaluations are consistent, fair, and

equitable (Howard & Gullickson, 2009). In other words, being systematic and documenting what is done with policies and procedures demonstrates that the process is done legally and ethically and for the welfare of the participants, which is an indispensable part of accuracy. In the case of the present study, the lack of policies and procedures, the lack of access to evaluation results, and the lack of a balanced feedback process in which only deficiencies are conveyed, negatively affected the accuracy of this case.

These current problematic situations in the school where this research was conducted are very similar to the findings of the research conducted by Collins (1999) in a private high school. Collins found that there is no written document to explain the purpose of the classroom evaluations. Furthermore, it was found that performance evaluation does not have a systematic process. Fowler (2001) conducted a study to analyze teachers' perceptions of their first-year experience in a growth-oriented teacher evaluation program. In this study, teachers suggested that documentation is needed, the summative assessment process should be explained better, and the evaluation process should be implemented more systematically. Teachers also stated that only in this way they can trust and fully implement the evaluation process. Some other studies also revealed that teacher evaluation is not systematic in many schools and even the purpose of the evaluation is not clear, or the teachers are not aware of this purpose (Fowler, 2001; Türkoğlu, 2015; Weisberg et al., 2009).

It is one of the important findings of the present study that the teachers were given only summative feedback in general meetings. It was also revealed that teachers did not receive any formative, constructive, or individual feedback about the effectiveness of their teaching. This situation experienced by teachers in the process of receiving feedback is also similar to the results of different studies. As in this school, which is the case of the present study, Collins (1999) also found that the process of feedback includes similar problems like making no explanation after the observations, only giving feedback about the negative situations, and giving feedback in general staff meetings. Research also showed that lack of detailed feedback (Sinnema & Robinson, 2007) and lack of effective formative feedback (Kraft & Gilmour 2016; Lavigne,

2014) are some of the problems faced by teachers. On the other hand, Widget Effect Report results showed that most of the teachers are not receiving specific feedback after evaluations to improve themselves which makes them feel like they are being treated in an injustice way (Weisberg et al., 2009). Furthermore, only one-quarter of fifteen thousand teachers reported that they have participated in a single informal conversation with the principal for over a year about improving instructional performance (Weisberg et al., 2009).

The absence of clear criteria and a functional reporting process for sharing evaluation results and feedback in the school, where the research was conducted, adversely affected both the usefulness of the evaluations and the defensibility of the results for the evaluators. In the process of making the results of evaluation useful for the teacher being evaluated, it is seen that the teachers still do not get enough support on how to use these results and they do not provide sufficient guidance about the feedback (Ford et al., 2016). Evaluation systems should summarize performance expectations using well-defined and clear criteria for the use of results (Howard & Gullickson, 2009). In addition, evaluations should use well-defined criteria to interpret or judge the performance based on a clear and defensible rationale. The lack of clear, precise criteria or standards is also criticized in existing teacher evaluation processes and many studies have demonstrated the importance of including certain standards in teacher evaluation systems (Collins, 2009; Donahue, 2016; Ilgaz, 2011; Lillejord et al., 2018; Kimball, 2001; Kraft & Gilmour, 2017; Royal & Tossman, 2009).

In addition, the existence of an environment that supports competition among teachers in the school also negatively affected accuracy. Teachers are compared with the average scores of students in school-wide exams. The teacher who teaches in the classroom with students with a high average score is appreciated more than other teachers. In other words, judgments are made about the effectiveness of the teacher according to the average success of the students in the school-wide exams. When this situation is examined, even if it is not written or explicitly stated it can be said that teacher evaluation is made to increase student success. Increasing student success is of course important, but in this school, the process of obtaining valid data was

negatively affected, as the evaluation of the teacher only with the scores of the students or only by checking over the lesson plans. This process of evaluation was not reflecting the complexity of teaching and how the teacher supports the development of students in other areas such as the affective domain. In such a school environment, teachers constantly prepare their students for school-wide exams and compete with other teachers, even if they do not want to. In general, teachers stated that they are not satisfied with teaching just for the sake of observation or being held responsible for student success. Different research results also reveal the similar problematic situations faced by the participants in this research for such outdated evaluations. Teacher evaluation systems that rely heavily on students' test score data may not reflect the effectiveness of the teacher but may also cause problems such as demoralizing teachers and encouraging effective teachers to leave the profession (Baker et al., 2010). Teacher evaluation with student achievement scores, namely growth scores, creates stress on teachers (Ford et al. 2016) and causes teachers to question whether the feedback given on these scores is defined fairly and accurately, rather than how to use this feedback in their development. Furthermore, teachers felt that the achievement scores provided insufficient detail about certain weaknesses and strengths of their teaching and were therefore undecided about what they should change in their teaching. Studies conducted with primary school teachers have also revealed that making biased evaluations, having problems in communication, and using the evaluation results for purposes such as comparing teachers, creating a competitive environment, and holding them responsible for student failure are problems that can be experienced in the teacher evaluation process (Süzen, 2007). Donahue (2016) talked about the similar problems that teachers encounter with teacher evaluation practices and mentioned that these problems create resistance for teachers to willingly adopt the model.

Decisions made in favor of experienced teachers and individual relations between teachers and teachers or between teachers and principals also affected accuracy as they cause biased evaluations. This privileged situation, which is given to experienced teachers, causes teachers to develop a belief that there is an unfair evaluation.

According to the “Widget Effect” report in 2009, it was revealed that experienced teachers were never fired for poor teaching performance and that the process of addressing poor teaching performance, which is the most basic assessment function of performance appraisal, was unsuccessful. Bernardin and Beatty (1984) stated that the participants ignored the feedback when they felt that the performance appraisal system was unfair, and the reliability of the feedback and sources was doubtful.

Since this research aimed to develop a teacher evaluation model aiming at professional development, the opinions of the participants were also taken about the professional development studies carried out at the school. It has been revealed that in-service training seminars in the school are carried out based on the topics chosen by most of the teachers or on popular topics. This training is usually given to all teachers collectively and the teacher rarely participates actively. General training given to the large masses is interpreted as a problem being ineffective since they do not meet individual needs, what is learned is not permanent, and the participants cannot participate. Teachers are generally not satisfied with in-service training (Karasolak et al., 2012; Göksoy, 2014; Özbek & Taneri, 2019). As research has proven, general one-size-fits-all training delivered in short courses or one-shot workshops is not sufficient to meet teacher professional development needs; research has revealed it should be balanced with professional collaborative learning and individual development plans (CDE, 2015). Professional development activities should be based on practice and organized according to the individual needs of the teachers (Özbek & Taneri, 2019). It can be said that in these professional development studies carried out in the case, traditional professional development paths, which are criticized as a top-down, one-size-fits-all approach, see teachers as passive recipients of knowledge rather than active agents involved in the construction of professional learning, are used (Lieberman & Miller, 2014; Little, 2003; Wei, 2009). Analyzes have shown that this traditional model of professional development is less effective in improving teachers' practice than sustainable, school-based models where teachers link learning to active ongoing practice (Kenedy, 2016; Wei, 2009).

5.2 Teacher Evaluation Model Needed

Purpose of The Evaluation Model

The purpose of this teacher evaluation model was determined "to evaluate teacher qualifications and to provide systematic support to teacher professional development depending on the evaluation process". Teaching is a profession dedicated to the development and benefit of students. For this purpose, teachers are lifelong learners who continuously reflect on their progress and are committed to improving teaching (Graham et al., 2015; Stronge, 2018). According to TALIS 2018 results embedding professional development as an integral part of the work of the teachers as illustrated by teachers getting 100 hours of professional development per year in Singapore (OECD, 2020b).

The importance of accountability, that is, measuring teacher success with student test scores, gained value, and less attention was paid to the formative challenge of using teacher assessment to improve teachers' teaching (Firestone & Donaldson, 2019). This oversight is notable because while the formative aspects of assessment may benefit the majority of teachers, accountability mechanisms are likely to apply to only a minority of teachers (Donaldson & Papay, 2015). While all teacher evaluation systems currently adopted and implemented include at least one other indicator or measure of teacher effectiveness (ie, teachers' systemic classroom observations), unfortunately in many states VAM scores are the primary focus. Models that provide accountability and use measures of student growth as the sole source for determining teacher effectiveness fail to recognize the needs of teachers as learners because, while these results may be thought to be certain, what educational philosophers and researchers have known about education for decades (teaching and learning is inherently ambiguous and complex) directly contradicts (Ford & Hewitt, 2020). It is hoped that the efficiency of teaching will be increased with the professional development of the teacher. However, preferring the incentives used to increase the quality of the teacher, not the teaching, and punishing the unsuccessful teacher is deceptive and even fatal for educational reforms all over the world (Fullan, 2011). On the other hand, with the

implementation of the "Every Student Succeeds Act (ESSA)" in the USA, the use of "growth models" or "value-added models", in which the success of teachers is measured by student test results, has decreased and schools have sought evaluations that support teacher professional development (Close et al., 2020). Considering the teacher evaluation trends in the United States in recent years and the aim of teacher evaluation in other countries, which are considered successful, it can be said that the aim of the teacher evaluation model designed specifically for the case of the present study coincides with the new trends of teacher evaluation models around the world. Furthermore, when the doctoral dissertations and research conducted in the field of teacher evaluation are examined, it is seen that the findings revealed the necessity and importance of evaluation models that support the professional development of the teacher (Bige, 2014; Evans, 2019; Fowler, 2001; Kimball, 2001; La Masa, 2005; Marzano, 2012; Moss, 2015; Nilsen, 2006; Taylor & Tyler, 2012; Süzen, 2007; Zarro, 2005).

Although the importance of teacher professional development has been emphasized for many years, this situation differs slightly for Turkish teachers in the results of the TALIS 2018 report. A rate of 71.8 percent of teachers in Turkey mentioned that professional development activities, have positive effects on their teaching, but this rate remains below the OECD average (OECD, 2020b). Furthermore, Turkey is located in the lowest six positions among countries in which teachers believe that professional development activities have a positive effect on their teaching. The reason teachers in Turkey cannot believe enough in the benefits of professional development may be related to how professional development practices are determined, for what purpose, and/or how they are carried out. In the school where the research was carried out, the teachers are faced with the problem that the general education given to large masses does not meet the individual needs, what is learned is not meaningful and the participants cannot actively participate in the professional learning process.

When the participants were asked for solutions and suggestions, they stated that these evaluation results will also guide the in-service training activities in the school since they will reveal their own needs. This finding shows similarities with the TALIS 2018

data. TALIS 2018 results revealed that teachers in Turkey believed that the in-service training that is carried out considering the prior knowledge and needs of the teacher is the most effective. Research also showed that teachers can develop professionally in schools where the needs of teachers are determined correctly (Icel, 2008; Jiang et al., 2015; Kini & Podolsky, 2016; Kraft & Papay, 2014).

The participants of the research were asked the question of which qualification areas should be evaluated and improved in line with the purpose they determined. It can be said that the emerging qualification areas (planning and preparation, instruction, reflective thinking, communication and collaboration, and professional responsibilities) overlap with the areas in the leading teacher evaluation models (Clayton, 2017; Danielson, 2013; Marzano & Toth, 2013; Marzano ve Simms, 2014; Shinkfield & Stufflebeam, 1997; Sloat et al., 2017; Toch & Rothman, 2008; Bender, 2005; Danielson, 2007; Danielson ve McGreal, 2000; Deiro, 2005; Egan, 2010; Geng et.al, 2019; Graham et al., 2015; Graham & Berman, 2018). While the participants stated that these areas should be evaluated, they also emphasized that the teacher's weaknesses in this area could be identified, and the strengths of the teachers should be revealed for balanced evaluation.

Evaluation Criteria

The criteria for each qualification area were included in the model and these criteria were also used in the model tools. Moreover, attention was paid to ensure that the criteria representing teacher behaviors are measurable and observable. According to the findings of the present study, the participants explained the importance of evaluation with certain criteria in line with two needs. One of them is the teacher's need to know which behaviors are evaluated according to what kind of criteria while being evaluated. As stated by Darling-Hammond (2012), while determining criteria for teacher evaluation, it is important to consider what behaviors teachers will be included in detail. For the developed teacher evaluation model not to leave out many important teacher behaviors and strategies, the list of criteria is kept very detailed in many subcategories.

Another need of the participants to use detailed criteria is the need-to-know what behaviors the teachers lack so that they can express themselves or defend themselves when necessary, during the process of explaining the evaluation results and giving feedback to the teacher. It has been determined that it is important for the accuracy of the model to justify the results by sharing the decisions obtained, to be meticulous about the results, and to pay attention to the fact that all results are based on justification criteria. Looking at the overall evaluation system, clearly defined performance standards at the beginning of the process, clear feedback based on multiple observations, and the performance rubric create a transparent system. This transparency can also build teacher enthusiasm for the process, foster more trusting and respectful relationships between teachers and evaluators, and increase the likelihood that the process will have an impact on teaching practices (Donahue & Vogel, 2018).

One of the challenges inherent in any rating system is the consistency or reliability of ratings given by raters. To increase consistency in teacher evaluations, detailed, standards-based performance rubrics should be adopted to provide objectivity and consistency among raters (Donahue & Vogel, 2018). Open standards or criteria can also make it easier for evaluation to be perceived as authentic and fair (Lavigne, 2014). According to Danielson and McGreal (2000), each teacher qualification should be represented in detail by criteria or standards to accurately determine the behaviors to be evaluated with teacher evaluation models, and valid indicators of the teaching approach that change over time can be added to these criteria. The participants of the present study also stated that inclusion of the criteria into an evaluation model would clarify what will be evaluated, that they can manage the acceptance or self-defense processes of the decisions taken according to the criteria, and in this way, they would trust the model. These findings, which are also consistent with the research findings conducted by İçel (2008), showed that well-prepared rubrics containing clear criteria or standards can be used to build trust between teachers and principals. Teachers generally liked rubrics containing detailed criteria and used them to reflect or evaluate their practices (Garet et al., 2017). Other studies have shown that data tools prepared

with clear criteria provide important feedback to teachers and inform the direction of professional development (Daley & Kim, 2010; Holtzapple, 2003; Kaplan, 2019; Özbek & Taneri, 2019).

Effective Classroom Observation

The teachers, principals, and experts participating in the research clearly stated that the evaluations made in this school do not reflect the real situation in the classroom and that what is happening in the classroom must be observed. For this reason, in this model, classroom observations were used as the basic evaluation tool. Similarly, the findings of Özbek and Taneri's (2019) study conducted with 304 principals and teachers; most of the participants stated that teacher performance evaluation can be done through classroom observation. Most studies investigating how observation affects teaching have found that observation can improve teaching or at least help teachers improve their teaching, even though the changes are small and uneven. According to the study conducted by Firestone and Donaldson (2019) observations can increase teachers' access to data and those teachers are likely to find these data formats more useful than growth scores. Teachers reported that observation data, unlike growth scores, can encourage productive conversations with colleagues and mentors about teaching, and in some cases, change their teaching (Firestone & Donaldson, 2019). Furthermore, it was revealed that classroom teachers, who are observed more in the classroom, find the observation process useful (Donaldson et al., 2014).

The fact that classroom observation provides data to the teacher and this data is meaningful for the teacher has led to an increase in the use of observation over the years (Firestone et al., 2014). Classroom observations are effective assessment tools and today's teachers often find data from classroom observations useful, as opposed to growth measures (Firestone & Donaldson, 2019). On the other hand, classroom observations can also bring some problems. It is common practice to use classroom observations to evaluate teachers, but there is much inconsistency between what to

focus on when making classroom observations, duration, and frequency of observation (Pianta & Hamre, 2015).

Observation may also cause some unexpected results, such as the fact that the observation creates time pressure, especially for the principals (Stecher et al., 2018), or causes the routine organizational tasks to be sacrificed because it takes the principals' time too much (Donaldson et al., 2014; Firestone et al., 2014). On the other hand, researchers identified reliability and validity issues with both observational and test-based measures of teacher performance (Baker et al., 2010; Kimball & Milanowski 2009; Lavigne, 2014). To prevent these negative aspects of the observation process, principals and school districts need to use classroom observation tools with proven reliability and validity as part of the teacher evaluation process, to ensure evidence-based teacher evaluation policy and practice (Marx, 2014; Pianta, 2012).

Determining whether the information produced as a result of the evaluation is valid enough to make decisions and make judgments affect accuracy (Howard & Gullickson, 2009). Misinterpretations should be avoided to ensure accuracy. Making judgments about teacher performance with a single observation in this school where the research was conducted affected the accuracy of the results. The teachers stated that they behaved differently from normal due to the anxiety they felt and the uncertainty of not knowing what the result of this observation would be used for, and this situation negatively affected the process of reaching the right results. Collins (1999) also revealed that observing the teachers with a single classroom observation for a year at the school where he conducted his research led to similar results. Furthermore, effective teaching depends on the type of the activity, the subject matter, pupil background, and pupils' characteristics (Jones et al., 2006) and it is not possible to observe all these variables together with a single class observation. Sufficient time for effective classroom observation and observations should be repeated many times. According to Widget Effect Report results' schools are having problems in identifying the poor or effective teachers although school members especially teachers and administrators both recognize ineffective teaching in their school due to short and

infrequent observations mostly conducted by principals without getting training (Weisberg et al., 2009).

For ensuring accuracy is that it is important in increasing the reliability of the data collected about the teacher due to the information gathered from more than one source. Therefore, when observing the effectiveness of a classroom teacher, the performance of the teacher in different lessons, different subject matter, and different course sections (teaching a new topic, reinforcing a topic, evaluating, etc.) should be evaluated by more than one person. This situation, which emerged concerning multiple observations, is like the results of the study conducted by Marzano in 2012. Throughout the study, many of the participants stated that a single observation made for one year is not enough and that the different course processes of the teacher should be evaluated by different people such as teachers and principals. Teachers and principals will contribute to the design and implementation of fair and transparent evaluation systems with multiple effectiveness measures (McQueen, 2022). In many research and literature sources teachers request multiple classroom observations to review professional practices (Cohen & Goldhaber, 2016; Collins, 1999; Ford & Hewitt, 2020; Fulton, 2019; Ilgaz, 2011; Jiang et al., 2015; Kaplan, 2019; Kimball, 2001; Winslow, 2015; Stronge, 2006; Süzen, 2007; Yılmaz, 2017).

It is estimated that the observation forms and the planned observation process in the evaluation model designed within the scope of the present study will prevent the mentioned problematic situations. The developed observation forms include comprehensive and detailed criteria for the teaching. Pre-pilot studies of these forms were carried out and necessary arrangements were made for their validity. It is estimated that the training based on the practical application will be given to those who will use the form, which would improve the competencies of the evaluators. The fact that teachers will evaluate multiple and repetitive measurements by making more than one observation makes the forms developed within the scope of this model powerful. Within the scope of this model, before the observation, relevant information will be collected about the lesson process to be observed through the interviews with the teachers. The lesson is observed in line with this information. Thus, the teacher

was allowed to express himself through the course process and to prevent any misunderstandings that may occur. For standardization, it is important to include rules, and procedures for data collection procedures to ensure consistency and quality in the data collection process (Pianta & Hambre, 2015). In the present study at the beginning of each tool used in the model, information, and procedure about how to use the tool and important points are given in detail. On the other hand, the length of the observation, training protocol, and scoring directions are kept standardized in the data collection process. Collins (1999) also draws attention to the conclusions regarding the importance and necessity of clearly specifying the rules and procedure of the data collection process.

Self-Assessment

The teachers and the principals working at the school where the research was conducted mentioned the importance of a classroom teacher knowing her teaching better than anyone, and the importance of evaluating herself correctly and objectively. Teachers know very well what kind of teacher they want to be by being aware of their personalities, abilities, and what they enjoy doing in the profession (Graham et al., 2015). According to Graham, Berman, and Bellert (2015), "failure to implement a plan means making a new plan for failure". Good lessons do not happen by chance, effective teachers apply what they have designed for teaching, question this practice, and use the results of inquiry (Marzano & Toth, 2013). Self-evaluation should be one of the data sources that should be included in performance appraisal to help teachers see their shortcomings and needs (Özbek & Taneri, 2019). Teachers need to think about lesson plans, teaching, assessments, and decisions and focus on the impact these elements have on student learning. In this way, teachers can decide whether their efforts require a change or to continue their effective practice in the future. If teachers can identify the gaps in their learning, that is, if they are aware of their strengths and weaknesses by questioning their teaching process, they can develop professionally (Hernandez & Endo, 2017). Within the scope of the model developed in line with the present study, teachers evaluate both their lesson plans and their teaching. The teacher is also asked to evaluate himself in the post conferences held after the classroom

observations, and he evaluates himself together with the school principal at the professional development meetings held at the end of each term. In short, within the scope of this model, the teacher is given the opportunity for self-evaluation and reflective thinking many times. Fulton (2019) also revealed that teachers were aware of the importance of evaluation components, and they perceived teacher reflection as the most influential component of improving teacher practices. Another research finding showed that teachers found the self-reflection process useful to support professional development and building teacher effectiveness over time (Jaffur, 2017). Along with this research, many studies reveal that the role of teacher self-assessment is important in teacher evaluation systems (Lillejord et al., 2018; Nelson, 2015; Yılmaz, 2017; Kaplan, 2019).

Lesson Plan Evaluation

The field of planning and preparation for the lesson includes planning and preparing the teaching process, materials, technologies to be used, and assessment processes of the lesson. The most effective teachers are also those who can plan well and organize processes (Stronge, 2018). The more effectively the teacher plans and prepares the lesson, the more likely it is to use instructional strategies effectively (Marzano & Toth, 2013). On the other hand, lesson plans give the observer detailed information about the focal point to be observed in the lesson, and an effective teacher is expected to carefully plan the teaching process for each lesson (Marshall, 2013). In the model prepared in line with the findings of the present study, the lesson plans of the teachers are evaluated with the help of a form consisting of detailed criteria.

Evaluating Communication and Collaboration

In the model prepared in line with the findings of the present study, communication and cooperation are evaluated by the head of the department and the classroom teachers with an evaluation form. In this form, positive communication and collaborative working areas are discussed. These areas were revealed in detail with criteria such as using positive language in communication, listening carefully to the other person, sharing teaching materials with colleagues, being open to learning from

colleagues, and taking an active role in group work. It can be said that the teachers of successful countries are in constant communication and cooperation with each other. An important factor that distinguishes Canada, Singapore, and Finland from many systems is a culture of collaborative professionalism that serves the teacher for both individual and collective learning (Fullan & Hargreaves, 2016). In addition to having very effective teachers and principals who perform very well in schools, effective practices should be shared in collaboration and supported by other teachers and principals. In other words, schools that constantly complain about their colleagues and that have a culture of competition among colleagues cannot make such shares, so it may not be possible to make progress (Fullan & Hargreaves, 2016).

In fact, at the core of the success of school systems is a culture of interaction, interactive pedagogy, mutual trust, and regular and quality feedback, which is a function of purposeful cooperation. (Fullan & Hargreaves, 2016). Teachers learn a lot from each other, and the effectiveness of their communication with each other in this professional learning environment also affects the quality of sharing and contributions (White, 2016). It is important to conduct communication as positively as possible, as people are generally more receptive and cooperative when approached in a kind, non-threatening manner (Bender, 2005). When you start the communication process with positive comments, this positive start continues throughout the communication process and helps create a more moderate environment when you need to address unpleasant but necessary issues. Another important aspect of communication is to use active listening strategies. Listening carefully to the other person, making you feel that you are listening to understand while listening, and using expressions that show you empathize show that you are listening effectively (Bender, 2005).

Professional collaboration among teachers in the school is an example of the correct execution of quality processes (OECD, 2018). Teachers work in collaboration with their colleagues to develop teaching processes, lesson plans, course materials, assessment processes, teaching approaches, strategies, activities, etc. They can share, exchange views, and make editions. Teaching is quite complex, and it may not be possible for every teacher to experience problems and solutions related to every

situation in the classroom or to use every teaching method and technique effectively (Danielson & McGreal, 2000). Each teacher needs to share unique application examples from their classroom and creative and effective solutions to the problems experienced in supporting each other's professional development, and the development of a school with collaborative teachers and the increase in job satisfaction and student success in this school will be inevitable (Graham et al., 2015; OECD, 2014). Teachers are mostly only listeners in the training they attend to support their professional development, however, when they work in cooperation with their colleagues, they learn from each other and are willing to practice; they feel part of the school community and learn more effectively than many in-service training programs (Lieberman, 2000; TEDMEM, 2018).

Evaluating Professional Responsibilities

Teachers working in a school spend most of their working hours attending classes and fulfilling their responsibilities to their students. At the same time, teachers try to improve themselves professionally outside the classroom to fulfill their other responsibilities at school. In today's world, students' profiles and educational innovations are rapidly changing. Cognitive learning processes, the functionality of learning, the meaning of learning, and how high-level thinking should be taught are the subjects that should be developed in every period of the teaching profession (Graham et al., 2015). It should not be forgotten that students' sustainable skills, that is, they can continue to learn and adapt to new conditions throughout their lives by taking advantage of students' learning abilities, can be also developed by teachers (Graham et al., 2015; Graham. & Berman, 2018). For teachers to develop these skills in students, they need to be open and willing to learn, follow the innovations, regularly update the knowledge they have, work with their colleagues for developmental purposes, participate in training, and read and search. Thus, it is an undeniable fact that professional development is a never-ending process for teachers. In another word, continuous professional development is an indispensable part of effective teaching, and this development brings student success.

It was very important for the case of the present study to evaluate the service to school including professional and in-school responsibilities of teachers such as making contributions to the development of the school, participating in the activities carried out in the school, keeping duties in the school, and taking responsibility in ceremonies. Teachers' fulfillment of their in-school responsibilities is very important in terms of the efficiency of the school operation and the sustainability of the school culture. Teachers stated that not everyone is equally involved in the fulfillment of these responsibilities, in this case, no sanctions are applied and therefore they evaluate the school environment as unfair. This sense of injustice seems to have made it important for teachers to measure this area. If it is important for a school that teachers fulfill their responsibilities outside the classroom, the teacher evaluation system may ask teachers to take responsibilities outside the classroom and contributing to the development of the school can be among effective teacher behaviors (Danielson & McGreal, 2000; Jenkin & Lord, 2006). In the model prepared in line with the findings of the present study, professional responsibilities are evaluated by the head of the department and the classroom teachers with an evaluation form, the criteria of which are prepared in detail.

Feedback Process and Functional Reporting

Many studies revealed that feedback is important in supporting professional development (Collins, 1999; Lillejord et al., 2018; Türkoğlu, 2015). To maintain progress with professional development efforts, evaluators should continually involve teachers in regular conversations with specific feedback (The New Teacher Project, 2010). When a teacher receives effective and formative feedback, their confidence in the evaluation process increases, and they think that it benefits their professional development. Overall, a large majority of teachers from TALIS countries (83.2% on average across TALIS countries) who had received appraisal and feedback considered them to be fair assessments of their work, and most of them (78.6%) found that they helped develop their work as teachers (OECD, 2009a). In the teacher evaluation model developed within the scope of the present study, feedback is given both through interviews with teachers after each observation and with an explanatory and functional

report at the end of each term. Immediately after the lesson observation, the evaluator conducts interviews with the teacher. In these interviews, the teacher is expected to evaluate himself and the teacher is allowed to explain the specific events or situations that happened in the classroom necessary. After the teacher's self-evaluation, the evaluator gives his feedback on how the teacher can improve herself in the necessary areas, with the help of criteria. Furthermore, in these interviews, teachers' opinions are also taken about how they can change or improve the parts that they think are missing or ineffective during the lesson.

Many studies have shown that effective feedback should be clear, detailed, useful, immediate, and evidence-based (Ford et al., 2016; Murtagh, 2014; Reddy et al., 2016). Feedback should include evidence of instructional practice or improvement goals for teachers (Hattie & Timperley, 2007; Looney, 2011; Tuytens & Devos, 2014). Teachers, who find their observational feedback generally helpful, believe that this feedback is fair, objective, or accurate, that they contain actionable suggestions for their improvement, and that they can have useful discussions on these suggestions (Garet et al., 2017; Jiang et al., 2015; Sporte et al., 2013). In the present study, post conferences were organized to have meaningful discussions for improvement, to talk about the suggestions, and thus give useful feedback to the teacher. Since it is generally thought that teachers cannot evaluate themselves objectively, this type of evaluation is avoided, but this also affects the active participation of the teacher in their evaluation process. Teachers believed that setting attainable goals during post-conference with their evaluators helped improve their teaching effectiveness (Vandermolen & Mey-Looze, 2021). Observation and the post-conference process gave teachers information that could improve how they teach (Sporte et al., 2013; Stecher et al., 2018). A study conducted by Jiang et al. (2015) also revealed that teachers agreed that the feedback they were provided in the post-observation conference guided them to improve their teaching. Similarly, another study finding showed that the teachers embraced post-conference as the most impactful part of the entire evaluation system (Jaffurs, 2017).

The reporting part of the model developed in line with the findings of the present study also provides detailed and specific feedback to teachers. In addition to the feedback process after each classroom observation results were gathered from lesson plan evaluations, all the classroom observations conducted throughout a semester, self-evaluation reports, professional development evaluation results, and communication and cooperation evaluation results will be reported at the end of each term. The holistic report, which is prepared based on the criteria in the forms, is shared with the teacher by the school principal through face-to-face meetings at the end of each semester. The meeting content is only accessible to the evaluated teacher and is kept completely confidential. In this "Performance Evaluation Meeting", teachers always have the right to express their opinions about the evaluation results. The main purpose of this meeting is to appreciate the teachers' identified strengths and to create a joint work plan to ensure that all other teachers benefit from these aspects. The second purpose of this meeting is to plan what can be done in a semester regarding the identified and agreed-upon weaknesses of the teacher and thus to support the professional development of the teacher. The specific feedback given to the teacher with the help of this reporting is directed towards both strengths and weaknesses. In addition, professional development suggestions will be useful to improve teaching as they emerge as a result of a detailed evaluation and are developed by establishing open communication with the teacher. Overall, this report will illuminate the teacher's concerns or doubts about the evaluation, as well as provide practical feedback to the teacher through an open dialogue. It is not surprising that an important part of what motivates teachers to improve is the desire to see their students grow and develop. There is a lot of research that the feedback given to the teacher who is so motivated by the student's development should be aimed at improving the teacher's teaching (Ford & Hewitt, 2020; Lavigne, 2014; Jiang et al 2015). The ultimate goal of reducing power inequalities and re-centering teachers as key actors in the assessment system is to ensure that feedback is used not only for effectiveness judgments but also to improve teaching (Ford & Hewitt, 2020). As Özbek & Taneri, (2019) revealed, providing timely and sound information on performance evaluation can prevent the spread of unnecessary fear and false information. The benefits of this open

communication and collaborative feedback can be found in other research and literature. Teacher evaluation models are effective when meaningful feedback processes are carried out in cooperation (Icel, 2008; Jiang et al., 2015; Kini & Podolsky, 2016; Kraft & Papay, 2014). In the field of education, while giving feedback on strengths and weaknesses, specific teaching strategies used in observed lessons, content-related topics, or improvement suggestions for any of these can be given through open dialogues (Hattie & Timperley, 2007; Looney, 2011; Tuytens & Devos, 2014).

The reporting process that emerged as a result of the present study was planned to be shared face-to-face by the school principal. Although teachers, in general, find it more beneficial to receive feedback from their fellow teachers, the research conducted by Reinhorn et al. (2017) showed that teachers can also receive meaningful feedback from their principals. According to the findings, teachers stated that they were evaluated by their principals for up to ten observations a year and received face-to-face feedback immediately after these observations. Furthermore, the teachers who participated in the study found these conversations particularly meaningful and useful. Stecher et al. (2018) also found that, in conjunction with improvements in teacher observation, principals have better knowledge to plan professional development and teachers use observation, among other influences, to select professional development.

It was frequently emphasized that teachers should have confidence in the evaluation process in the school where the present study was conducted. Teachers have frequently reported that they need to trust this model to accept and apply assessment results. A trustworthy assessment environment is closely related to the feedback process. If the teacher believes that the feedback received by the principal is in their favor, and their sole purpose is to improve themselves, they trust this system. According to the study conducted by Donahue and Vogel (2018) teachers felt that feedback was more valuable when it came from a reliable and reputable source. Some of the teachers who took part in the research expressed their belief that if the relationship between themselves and their evaluators was not positive, there would be no change in their teaching practices. Increasing reliance on principal feedback can increase teachers'

willingness to use assessment feedback for self-improvement (Awkard, 2017). Once trust is established between the teacher and the administrator providing the feedback, the teacher is more likely to perceive the evaluation feedback as valuable in the observation process (Awkard, 2017; Yoo, 2016).

Objectivity

The opinion that teachers need to believe in the evaluation process and the objectivity of the evaluator is one of the important elements that shape this evaluation model. The participants described the comparison and competition-based environment created among the teachers, the different treatments made in favor of the experienced teachers, the use of student achievement as a single source in evaluating teacher effectiveness, and the neglect of the instruction and effort in the classroom as problems that hinder objectivity. They stated that this problem should be solved in the teacher evaluation model that was developed through the present study.

To ensure objectivity and accuracy of the model providing its validity orientations is one of the most important findings which comprises independent evaluations apart from prejudices and personal relationships. To achieve this, one of the important issues included in the model is the selection of the observer. Teachers will be evaluated by teachers from different grade levels in order not to affect the evaluation process due to the personal relations between observers and teachers. The fact that the classroom teachers teach at different grade levels each year they master all four levels, and this situation enables teachers working at different levels to evaluate each other. Furthermore, to ensure objectivity and the consistency of the evaluations, attention will be paid to the evaluation of the teacher's lessons by two different observers at the same time. Present study participants also stated that the evaluator's classroom teaching experience would provide more reliable data, which is in line with the findings of the research conducted by Donahue and Vogel (2018). Similarly, in this study, the teachers stated that if the evaluator has sufficient teaching experience, they can evaluate the teaching in the classroom correctly and give accurate feedback on what is what.

Another finding that is specific to this school and stands out in ensuring accuracy is that everyone should be evaluated equally and regardless of their teaching experience. This finding is specific to this school, and the main reason for this is that teachers are uncomfortable with different practices in favor of experienced teachers and the oppressive approach that experienced teachers take against new teachers. These practice differences among teachers undermine their trust in the school and evaluation system. For instance, in a study high school teachers stated that some teachers were treated as privileged, the warnings given to them were not given to these teachers and they considered this situation unfair (Türkoğlu, 2015). Through the developed model of the present study, regardless of new or experienced teachers, each teacher is evaluated with the same data collection. After these evaluations, planning will be made for the new or experienced teacher to receive training in any field that they need individually. Research from "Project on the Next Generation of Teachers" revealed that the training received by new teachers within the scope of professional development supported them and they felt safe in the profession thanks to the guidance they received (Johnson & Kardos, 2002). This result of the research supports the claim that this school-based model will be effective in meeting the professional needs of the teacher in the school.

Overall if the purpose of the evaluation is professional development, then a more trusting environment is needed between teachers and principals or other stakeholders. Teachers, who are judged by the growth scores obtained through student test results, feel increased pressure and competition, which reduces morale and cooperation (Collins, 2014). The study conducted to support efforts in increasing teacher effectiveness by examining how teachers value the feedback they receive showed that credibility was the most important characteristic affecting teachers' response to feedback (Cherasaro et al., 2016). In other words, according to this study, teachers' reactions to feedback are affected by how useful they perceive it, which in turn, is how credible they perceive their evaluators. As Özberk and Taneri (2019) found in their study, it will be effective if teacher evaluation is done accurately, unbiased, fairly, and scientifically. When we look at the present study in general, it seems more likely that

systems or models that are far from prejudice, where teachers are not competing with each other, and where objective and fair evaluations are provided, are more likely to be effective, as frequently stated by the participants of the case. It seems very important that the person being evaluated believes in the evaluation system and its effectiveness, as in many other evaluation systems. In the study conducted by McQuen (2022), teachers also talked about their feelings of anxiety, bias, and unfairness toward the evaluation system. The teachers in this study felt that the teacher evaluation system was effective when the methodology was objective. The results of the findings noted that most respondents wanted evaluations to be fair, consistent with open communication, and as objective as possible. It has been revealed that teachers want to believe that the results of high-stakes evaluations are fair and that they are transparent to the criteria used to evaluate teachers (Hallinger et al., 2014). Von der Embse et al. (2016) also stated that the teachers perceived the selected evaluation methods as ineffective, that they used ineffective methods such as fear tactics by the administrators in the assessments, and as a result, they generally experienced high-stress levels. This is in line with the findings of the McQueen (2022) study, as this research shows that the evaluation system and method play a role in their thinking about whether the evaluation system is subjective or objective.

Openness and Confidentiality

The teachers working in this school emphasized that feedback should be given face to face, open and positive communication should be established while giving feedback and a supportive approach should be followed. In addition, it has been revealed that it is important for the accuracy of the model that teachers have a say in the information collected about themselves, communicate with the evaluator, and defend themselves in a positive and open environment. Since observation is the basic data collection tool of this model, instant feedback is given to the teacher after the observation. Giving this feedback through positive communication along with the effectiveness of the content facilitates the teacher's process of accepting criticism. The implications of the findings of the Paufler et al. (2020) study showed that the teachers in this study find that the assessment system is beneficial when they are actively involved in the process

and have positive relations with the evaluator. According to this research, post-observation interviews allowed teachers to better understand and implement improvement suggestions because they were able to ask explanatory questions and specific guidance. If teacher evaluation systems are based on classroom observation, the teacher can be informed with more constructive discussion, and a standardized observation system helps to establish a common language for such conversations (Stecher et al.2018). According to Stronge (2018), teachers who work in a more supportive professional learning environment improve their effectiveness more quickly than those working less supportive context. Evaluation systems provide teacher improvement best when there is two-way communication and a supportive climate is established (Ford & Hewitt, 2020). In a case study conducted with high school teachers, it was concluded that teachers needed motivation, they considered principals to use positive communication and they expected supportive guidance from principals rather than commanding words (Türkoğlu, 2015). Other research also showed that teachers develop rapidly professionally when they receive positive, supportive, and face-to-face feedback and communicate positively with the evaluator (Hallinger et al., 2014; Kimball, 2001; La Masa, 2005; Winslow, 2015).

Dynamics-Evaluator Trainings

In the model developed within the scope of the present study, it is planned to provide training at regular intervals to those who will evaluate and be evaluated. In these training, practical courses will be given about the purpose and process of the evaluation model, the effective use of data collection tools, and the reporting process. Considering that observation is an important part of this evaluation model, examples such as videos taken in the classroom environment, transcripts are written for classroom observation, and hypothetical scenarios will be used in the training to be held before the model is applied. With these examples, practical implications are made by using observation forms and evaluators are guided to look at a lesson observation from the same point of view. The training is planned to be repeated at regular intervals every year and the content of the training will be updated in line with the suggestions received. Widget Effect report showed that school principals are ill-equipped to

evaluate teachers effectively and school districts invest in evaluation training of administrators minimally (Weisberg et al., 2009). Training of evaluators is needed for a clearer understanding by teachers and principals of what is a great teaching and how to evaluate effective teaching (Danielson & McGreal, 2000).

As in this research, classroom observation tools have taken on a greater role in research and practice in recent years because of their potential to provide a valid, reliable, and evidence-based measure of teaching quality that can help rationalize teachers' supervision and evaluation. Employees at the school where the present research was conducted suggested that ensuring consistency between observations could be possible by training the observer. According to White (2016), it depends on carefully trained and supervised observers (evaluators) to make observations accurate and provide consistent scores. According to Pianta and Hambre (2015), to implement tools for practitioners with fidelity within the workflow and sustain the roles successfully at school it is important to apply a training protocol to all the raters. Daghe (2018), examined different teacher evaluation models and it was revealed that evaluators were getting training, especially about the rubrics used in the evaluation model and participants believed that the training about the use of the evaluation rubric gave them better credibility and ability to understand what to look for during the observations. Furthermore, the importance and necessity of evaluators to be competent in the field of teacher evaluation have been frequently emphasized in the present study and it is an indispensable element for this exemplary situation. Similarly, according to the findings of another research that reveals the importance of training, classroom teachers are considered trained and competent evaluators, and clear and straightforward standards and adhered to procedures as the most critical and desirable components of a teacher evaluation system (Nelson, 2015). Along with this research, many studies reveal that it is important to include evaluator training in teacher assessment systems (Grissom & Youngs, 2015; Kane & Staiger, 2012; La Masa, 2005; OECD, 2009b; Özbek & Taneri, 2019).

5.3 A School-Based, Participatory Teacher Evaluation Model

School-Based Evaluation Model

One of the biggest mistakes applied in developing the teaching profession is to believe in the truth of the general and universal and to apply it as it is (Lewis & Hogan, 2016). Schools should be considered with their teachers, principals, students, and school culture, and teacher evaluation approaches focusing on professional development should be designed by the context and goals of the school (OECD, 2013a). Schools should be the main organizations that support teacher learning, and schools themselves should be thought of as “learning organizations” (Bautista, 2015).

Although all schools are part of the same system, the professional development needs and development plans of each school will be different due to the different environmental conditions they are in and the needs of students and teachers (TEDMEM, 2018). It seems easier and more convenient to take the model prepared for teacher evaluation for schools and apply it as it is, but these applications do not meet the needs of the school, do not show the expected effect, and the use of such models is abandoned over time (Ofsted, 2018). It is important for teacher evaluation reforms to change the organizational culture at the school level and to ensure the effective participation of principals and teachers for their effective implementation (Dee et al., 2021). According to TALIS 2018 results teachers report that school-based professional development based on collaboration has the most impactful effects on teaching practices and is more relevant to the daily jobs of participants (OECD, 2020b). Adopting school-embedded approaches to teacher training is an efficient way and respond to the needs of teachers. One of the best ways for schools to conduct their collaborative and school-based professional development is doing peer observations where teachers have the opportunity to observe new pedagogical methods, evaluate the instruction of peers, and provide valuable feedback to foster the reflective practice and improvement like lesson study model used in Japan (OECD, 2020b).

The evaluation model, developed in line with the present study, is specific only to the culture and dynamics of this school. This school-based model, which was developed,

also supported school-based professional development because it aimed to determine the needs of the teachers accurately and clearly at this school with a school-based teacher evaluation. Thus, professional development studies will also be planned to meet these needs. Most of the professional development of Singaporean primary and secondary teachers takes place in school settings, where they have numerous work-embedded learning opportunities (Bautista et al., 2015). Just as the feedback given to students is successful when associated with their learning environment, teacher evaluation will not work unless it is placed in a school culture where teachers are motivated to learn by receiving feedback (Fullan, 2011).

Participatory Development Process of the Model

The teacher evaluation model developed in this research was developed with the participation of stakeholders (classroom teachers, principals, experts) working at the school. Teachers especially care about having a say in the planning and implementation of teacher evaluation models. What is known about how evaluation experience can change teacher effort and effectiveness from teachers' perspectives is relatively limited (Tuytens & Devos, 2013; Taylor & Tyler, 2012). Although there is a limited number of studies conducted to understand teachers' perspectives for developing teacher evaluation systems, to make scratch the surface of teacher evaluation development processes Jiang et al. (2015) conducted a study. This study aimed to understand teachers' perceptions of evaluation, develop reliable measures of these perceptions, and understand teachers' evaluation experiences to be useful both for researchers and policymakers. Fowler (2001) found that one important suggestion of the teachers was about stakeholder involvement meaning that teachers being involved in the design of the teacher evaluation plan to be clear about the elements of the evaluation, understand and apply the rubric, to be clear on how the elements of the evaluation related with rubrics and summative evaluation. The teachers stated that not getting their views while planning and implementing this system created the feeling that they dictated something, and the management was challenging them.

Research and literature studies that emphasize the importance of incorporating and co-developing the views of teachers and other stakeholders in the development of the teacher evaluation system, in general, have shown that teacher evaluation systems will be more effective (ESSA,2019; McQueen, 2022; Paufler et al., 2020). Teachers, students, parents, administrators, and officials involved in the development of teacher evaluation programs understand the importance of teacher evaluation systems (ESSA, 2019). The creation of teacher evaluation systems that consider the perception of the teacher about the effectiveness of this system can help determine how effective this evaluation system can be in terms of student success and teacher professional development (McQueen, 2022). The importance of involving evaluators and teachers in the evaluation process has also been demonstrated in a study conducted in Texas (Paufler et al., 2020). The perceptions of the teachers and administrators in this study showed how the perceptions of the participants and how these affect the impact and sustainability of the evaluation system in practice. Thus, in theory, the purpose, criteria, processes, and impact of assessment systems should work together to help teachers develop, capture the real work teachers do, provide opportunities for teachers to participate actively in the process, and have a positive impact on teachers and their students (Paufler et al., 2020). Yilmaz (2017) examined the evaluation systems of the four countries (China, Singapore, South Korea, and Japan) that were successful in the PISA 2012 exam, and she revealed that all stakeholders had a say in the evaluation process and the purpose of these evaluations was to ensure continuous professional development. Receiving input and feedback from teachers about teacher evaluation for teacher development that directly impacts students can help formulate teacher evaluation systems that are effective for teachers (McQueen, 2022). In recent years, collaborative and participatory teacher evaluation models continue to be developed to improve professional practices in schools (Darling-Hammond, 2014; Lillejord et al., 2018; Shulha et al., 2015).

It can be said that the process of developing a participatory approach positively affects the feasibility of the model. In the present study, the teachers stated that they were willing to be evaluated with this model in which their opinions were reflected.

According to the qualitative study conducted by DiGrazia (2018), it was revealed that teachers want to have more say in the process in which they are evaluated to meet the needs of the teachers, to have more control over the evaluation process and to accept the evaluation willingly. According to the findings of the study conducted by La Masa (2005), teacher evaluation would be more productive in improving teacher practice if the teachers play an active role in the decision-making process of the professional development and evaluation processes. Research conducted in a school in which an evaluation model was developed based on the participation of the stakeholders and results revealed that school staff generally had a clear idea of why evaluation was important, they viewed the evaluation as a tool that could foster enthusiasm and encourage staff in school improvement activities and participants were feeling a commitment to the evolution process (Lee & Cousins, 1995).

5.4 Implications for Educational Practices

Policymakers, practitioners, and researchers have paid a lot of attention to teacher evaluation reform over the past decade. Nations, states, regions, and schools have invested substantial financial, political, and human capital into overhauling their assessment systems. Teacher evaluation—as a legal requirement—is and likely will remain an important policy tool in many countries. The teacher evaluation process has always been important from past to present, and it looks like it will maintain its importance for many years. In order to design effective teacher evaluation systems, more investment is needed in developing models that include formative assessments that focus on professional development, especially in Turkey.

Rather than suggesting a teacher evaluation model that can be used by every school, teacher, and principal, this research is important in that it guides how to design an effective teacher evaluation model for the culture of a school and the needs of teachers and principals working in this school. The idea on which the research is based is that it is not right to evaluate without involving the evaluator in the work of the development process of the evaluation. For this reason, teacher evaluation models to be developed for schools should be made with the participation of teachers, principals,

and other relevant people working in the school. Otherwise, no matter how effective external teachers' or top-down evaluation models are, teachers' biased responses may prevent the implementation of these models.

Lillejord et al. (2018), defines teacher evaluation as a wicked policy problem and teacher evaluation can be successfully implemented by productive approaches, not through traditional linear approaches. Although this research being a case study is not a complete solution to this wicked problem, the results revealed what kind of teacher evaluation the teachers, principals, and experts working in a school needed. When the results are examined, the evaluation model that emerges includes the definitions and elements that many researchers in this field have made under the name of "effective teacher evaluation systems". It can be said that the qualification areas included in the model developed with the help of the results obtained from this research can be included in almost every teacher evaluation model. According to the needs of other schools, the criteria included in these qualification areas may differ or a weighting can be made between the qualification areas, unlike this case study. Moreover, it is thought that the basic elements that make up this model (multiple assessment, giving functional feedback, being participatory, ensuring objectivity, consistency, clarity, evaluator training, etc.) will also guide the teacher evaluation models to be designed.

Teachers are not happy to be evaluated with student achievement scores and these evaluations impose sanctions on them. Moreover, they do not trust the teacher evaluation models that are implemented by taking student growth measures as the basis, do not take the feedback into account, and think that this process is not fair. Since the participants did not trust the school-wide exams applied in this case, they objected to gather data from student success for teacher evaluation. If the examination systems applied at the school are changed, the opinions of the participants may change. We can say that we follow the individual development of the student when the minimum skills expected to be acquired by the students at each grade level are defined and these skills are measured and reported in the process through exams, observations, or assignments, instead of one-time exams for certain achievements and only true and false feedback is given as in this school. In this case, perhaps the participants may

want to be evaluated with the help of their students' individual development results. Education systems need student assessment processes in which the individual development of students is followed reliably so that these results can be a resource for teacher development. Furthermore, teacher evaluation models or systems should support and enable teacher professional development that aim to evaluate the important qualifications of teachers with detailed criteria and professional development activities should be planned to eliminate the weaknesses determined in these criteria.

The most important element for the effective implementation of teacher evaluation is that the practitioners trust and internalize the model or system. An effective appraisal system that allows teachers to participate in the assessment process, should be fair and impartial, continuous, and based on reliable and valid criteria as well. In almost every part of the study, teachers frequently stated that they needed to trust the model. The answer to the question of “What kind of a teacher evaluation model does a teacher trust?” can be found based on the findings of this research.

Evaluation models or systems that the teacher will trust should include the teacher's views, the teacher's needs should be determined reliably from multiple sources, and the feedback should be functional and feasible. Based on the findings of the present study, it can be said that the participants are willing to apply the evaluation model that they believe, trust, and find useful. It can be said that the need for teachers to trust the evaluation system, which emerged from the beginning of the study, can be eliminated by designing a model that will accurately determine the needs of the people to be evaluated and meet this need with effective feedback. When teachers understand the necessity of an evaluation process and are allowed to participate in the design and development of assessments, their attitudes and perceptions can be very positive (Özberk & Taneri, 2019). If the teacher has a positive attitude towards positive evaluation, the teacher will be willing to accept constructive criticism to create instruction that will increase student success (Nelson, 2015). Similarly, authentic feedback from a trusted and respected source encourages self-reflection and increases the likelihood that a teacher will act on new knowledge and insights gained from the

process (Donahue & Vogel, 2018). Looking at these results if we want teacher evaluation to work for the teacher and improve his/her teaching, what this data contains and how it is shared is also important along with the collected data. The teacher evaluation models to be developed should enable the teacher to access the evaluation results more, the usefulness of these results for the teacher should be based on, discussion and mutual opinions should be made about these results while the results are reported. Moreover, the teacher evaluation system should allow for feedback and collaboration between the evaluator and the teacher.

Moreover, it should include organizing training to improve the competencies of the evaluators and to inform everyone to be evaluated. In other words, for teacher evaluation models or systems to be developed in schools, evaluators should be trained on how to effectively implement the teacher evaluation system used to evaluate teachers.

5.5 Implications for Further Research

While designing the present study, the researcher was aware of teachers' reactions to top-down evaluations. Moreover, the necessity of developing a model-specific for the school's culture, ecosystem, and the relations of the stakeholders with each other was revealed when the literature was searched. It can be said that this importance was repeatedly demonstrated by the participants while collecting data for the present study. Although there is a tendency to develop and use performance evaluation systems throughout the country and generally in which the teacher is evaluated by scoring certain items, teacher evaluation models must be autonomous and school specific. For this reason, it is very important that teacher evaluation models, which will be developed with the help of further research, have the necessary flexibility to enable the application in schools and, if possible, be specific to the school.

Although the school where this research was conducted was a private school, there was no systematic process for teacher evaluation at the school. It can be said that there are teacher evaluation systems that have just started to be implemented in many private schools or have been applied for many years. Future research can be conducted to

evaluate the effectiveness of teacher evaluation systems implemented in these schools and determine the working or non-functioning aspects of these systems, and even develop new systems for rearranging the non-functioning aspects.

In the recruitment process of this school, which represents the case of the present study, not many inexperienced teachers were preferred, and most of the teachers were experienced teachers. Moreover, less experienced teachers often emphasized that the same evaluation should be made for every teacher since there is a biased evaluation in favour of experienced teachers. This result, which emerged following the teacher profile of this school, could have emerged differently in institutions where less experienced teachers work. For this reason, it may be important to conduct future research, especially in institutions where inexperienced teachers work.

The model that emerged as a result of this research could not be implemented due to both the school's policies and the covid 19 epidemic. The implementation and revision of teacher evaluation models, which will be developed in line with further studies, will be much more effective for practitioners in this field. In addition, the contribution of the model that emerged as a result of this research on the professional development of the teacher and the effects of the contributions of the professional development of the teacher on student achievement could not be determined. More longitudinal studies can be conducted to examine how teacher evaluation affects the professional development of teachers and to investigate how the student is affected by the teacher's development as a result of the evaluation.

As demonstrated in this study, teachers need to receive effective feedback after evaluation. Examining policies regarding the usefulness of feedback or collecting data to identify potential barriers to providing effective feedback could be the subject of research for future studies. Since perceptions of evaluator credibility are strongly associated with teachers' perceptions of the usefulness of feedback, focusing on ways to build evaluator credibility may be a topic for further research.

Furthermore, it has been determined both by this research and many other studies that it is important for the evaluated persons and evaluators to receive training. Research

can be conducted to plan the content of these training. Thus, research can be designed to determine how these training can be organized based on practice, what type of training should be provided to validate assessments, and in particular, ways to strengthen the usefulness of feedback.



REFERENCES

- Ahmad, Z. (2020). Peer observation as a professional development intervention in EFL pedagogy. *International Linguistics Research*, 3(1), p1. <https://doi.org/10.30560/ilr.v3n1p1>
- Alay, G. (2006). *İlköğretim ve orta öğretim kurumlarında görev yapan öğretmenlerin performans değerlendirme sistemine ilişkin görüşleri* (Publication No. 191148) [Master's thesis, Gazi University]. YÖK Thesis Database.
- Almutairi, T., & Shraid, N. (2021). Teacher evaluation by different internal evaluators: Head of departments, teachers themselves, peers, and students. *International Journal of Evaluation and Research in Education (IJERE)*, 10(2), 588. <https://doi.org/10.11591/ijere.v10i2.20838>
- Amrein, A., & Berliner, D. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10, 18. <https://doi.org/10.14507/epaa.v10n18.2002>
- Amrein-Beardsley, A., & Geiger, T. (2017). All sizzle and no steak. *Phi Delta Kappan*, 99(2), 53-59. <https://doi.org/10.1177/0031721717734191>
- Amrein-Beardsley, A., & Geiger, T. (2022). The breakfast of champions: Student surveys to sustain and cultivate teachers. *Phi Delta Kappan*, 103(8), 49-53. <https://doi.org/10.1177/00317217221100011>
- Archer, J., Cantrell, S., Holtzman, S., Joe, J., Tocci, C., & Wood, J. (2016). *Better feedback for better teaching* (1st ed.). San Francisco: Bill & Melinda Gates Foundation-Jossey Bass.
- Attard, K. (2016). Personally, driven professional development: reflective self-study as a way for teachers to take control of their own professional development. *Teacher Development*, 21(1), 40-56. <https://doi.org/10.1080/13664530.2016.1218363>

- Awkard, T. (2017). The power of reflective action to build teacher efficacy. *Phi Delta Kappan*, 98(6), 53-57. <https://doi.org/10.1177/0031721717696479>
- Bader, G., & Rossi, C. (2002). *Focus groups: A step-by-step guide* (3rd ed.). The Bader Group.
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L., Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers* (Briefing Paper No. 278). Economic Policy Institute
- Bautista, A., Wong, J., & Gopinathan, S. (2015). Teacher professional development in Singapore: Depicting the landscape. *Psychology, Society, & Education*, 7(3), 311. <https://doi.org/10.25115/psye.v7i3.523>
- Bender, Y. (2005). *The tactful teacher*. Nomad Press.
- Bernardin, H., & Beatty, R. (1984). *Performance appraisal: Assessing human behavior at work*. Kent Publ. Co.
- Bige, E. (2014). *İlkokul müdürlerinin ders denetimleri ile ilgili öğretmen görüşleri* (Master's thesis). Adnan Menderes University.
- Bilir, M. (1991). *Türk eğitim sisteminde teftiş alt sisteminin yapı ve işleyişi* (Publication No. 16815) [Doctoral dissertation, Ankara University].
- Bill & Melinda Gates Foundation. (2012). *Gathering feedback for teaching combining high-quality observations with student surveys and achievement gains*. MET project publications.
- Bill & Melinda Gates Foundation. (2013). *Ensuring fair and reliable measures of effective teaching; Culminating findings from the met project's three-year study*. MET project publications.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. doi: 10.1080/0969595980050102

- Bogdan, R., & Biklen, S. (2007). *Qualitative research for education: An introduction to theory and methods* (5th ed.). Pearson.
- Bonner, S., Torres Rivera, C., & Chen, P. (2018). Standards and assessment: coherence from the teacher's perspective. *Educational Assessment, Evaluation and Accountability*, 30(1), 71-92. <https://doi.org/10.1007/s11092-017-9272-2>
- Bookhart, S. (2020). Feedback and measurement. In S. Brookhart & J. McMillan (Eds.), *Classroom Assessment and Educational Measurement* (1st ed., pp. 63-79). New York: Routledge.
- Boydak Özan, M., & Karagözoğlu, A. (2020). Öğretmen performans değerlendirme taslağı: Görüşler, yansımalar, değerlendirmeler. *Journal Of Anatolian Education Research*, 4, 21-45.
- Bozan, S., & Ekinci, A. (2019). Okul müdürlerinin öğretmen performans değerlendirme yeterliliklerinin okul müdürleri ve öğretmen görüşlerine göre değerlendirilmesi. *Elektronik Sosyal Bilimler Dergisi*, 18(69), 142-161. <https://doi.org/10.17755/esosder.412344>
- Brandenburg, R., Glasswell, K., Jones, M., & Ryan, J. (2017). *Reflective theory and practice in teacher education*. Springer.
- Brandon, J., Adams, P., Friesen, S., Hunter, D., Koh, K., Mombourquette, C., Parsons, D., & Stelmach, B. (2018). *Building, supporting, and assuring quality professional practice: A research study of teacher growth, supervision, and evaluation in Alberta*. A research report for Alberta Education, Edmonton, AB.
- Briesch, A., Volpe, R., & Floyd, R. (2018). *School-based observation: A practical guide to assessing student behavior (The Guilford practical intervention in the school's series)*. The Guilford Press.
- British Columbia Institute of Technology (BCIT), 2010, *Classroom Assessment Techniques: Instructional Job Aid*, Burnaby.
- Brophy, J. (1992). Probing the subtleties of subject matter teaching. *Educational Leadership*, 49, 4-8.

Brumback, G. (1988). Some ideas, issues and predictions about Performance Management. *Public Personnel Management*, 17(4), 387-402. doi: 10.1177/009102608801700404

Buluç, B. (1997). Türk eğitim sisteminde teftiş ve denetim alt sisteminin gelişim süreci. *Bilgi Çağında Eğitim*.

Bülbül, T., Tunç, B., Özdem, G., & İnandi, Y. (2013). *Okul temelli değerlendirme algısı: öğretmen görüşlerine dayalı betimsel bir analiz çalışması*. ESTP, 1-20.

Carbaugh, B., Robert, M., & Toth, M. (2020). *New to the Marzano teacher Evaluation Model? The Marzano focused teacher evaluation model: a focused, scientific-behavioral evaluation Model for Standards-Based Classroom*. Learning Sciences International.

Campbell, A. B. (2014). *Understanding the teacher performance evaluation process from the perspective of Jamaican public school* (Publication No. 3631163) [Doctoral dissertation, University of Toronto]. ProQuest Dissertations and Theses Global.

Centre for Development and Enterprise (2015). *Teacher Evaluation Lessons from other countries*. CDE Publishing.

Center for Educational Leadership, (2022, May 4). *5D+™ Rubric for Instructional Growth and Teacher Evaluation*. <https://k-12leadership.org/tools/5d-rubric-for-instructional-growth-and-teacher-evaluation/>

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.

Chen, H., Li, M., Ni, X., Zheng, Q., & Li, L. (2021). Teacher effectiveness and teacher growth from student ratings: An action research of school-based teacher evaluation. *Studies In Educational Evaluation*, 70, 101010. <https://doi.org/10.1016/j.stueduc.2021.101010>

- Cherasaro, T. L., Brodersen, R. M., Reale, M. L., & Yanoski, D. C. (2016). *Teachers' responses to feedback from evaluators: What feedback characteristics matter? (REL 2017–190)*. Regional Educational Laboratory Central.
- Clayton, C. (2017). Raising the Stakes: Objectifying Teaching in the edTPA and Danielson Rubrics. In J. Carter & H. Lochte (Eds.), *Teacher Performance Assessment and Accountability Reforms: The Impacts of edTPA on Teaching and Schools*. Palgrave Macmillan.
- Close, K., & Amrein-Beardsley, A. (2018). Learning from what doesn't work in teacher evaluation. *Phi Delta Kappan*, 100(1), 15-19.
<https://doi.org/10.1177/0031721718797115>
- Close, K., Amrein-Beardsley, A., & Collins, C. (2020). Putting teacher evaluation systems on the map: An overview of states' teacher evaluation systems post-every student succeeds act. *Education Policy Analysis Archives*, 28, 58.
<https://doi.org/10.14507/epaa.28.5252>
- Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics Of Education Review*, 26(6), 673-682.
<http://dx.doi.org/10.1016/j.econedurev.2007.10.002>
- Cohen, J., & Godhaber, D. (2015). Observations on evaluating teacher performance: Assessing strengths and weaknesses of classroom observation and value-added measures. In J. Grissom & P. Youngs (Eds.), *Improving Teacher Evaluation Systems: Making the Most of Multiple Measures* (pp. 8-22). Teachers College Press.
- Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher*, 45(6), 378-387. <https://doi.org/10.3102/0013189x16659442>
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. Routledge.
- Cochran-Smith, M. (2019). Teacher education for justice and equity: 40 years of advocacy. *Action In Teacher Education*, 42(1), 49-59.
<https://doi.org/10.1080/01626620.2019.1702120>

Collins, A. (1999). *A case study of instructional supervision system at a private secondary school* (Publication No. 82098) [Doctoral dissertation, Middle East Technical University]. YÖK Thesis Database.

Connell, R. (2009). Good teachers on dangerous ground: Towards a new view of teacher quality and professionalism. *Critical Studies in Education*, 50(3), 213–229.

Cousins J. & Earl L. (1995). The Case for Participatory Evaluation: Theory, Research, Practice. In J. Cousins & L. Earl (Eds.), *Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning*. Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning. (pp. 1-21). Routledge.

Creswell, J. (2013). *Qualitative inquiry & research design* (3rd ed.). Sage Publications.

Creswell, J. (2015). *30 Essential skills for the qualitative researcher*. Sage.

Creswell, J., & Poth, C. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (4th ed.). Sage.

Culbertson, J. (2012). Putting the value in teacher evaluation. *Phi Delta Kappan*, 94(3), 14-18. <http://dx.doi.org/10.1177/003172171209400304>

Çavuş, A. (2010). *Ortaöğretim kurumlarında görev yapan yönetici, öğretmen, öğrenci ve velilerin çoklu değişkenlere göre performans değerlendirmeye ilişkin görüşleri* (Publication No. 278137) [Master's thesis, Gazi University]. YÖK Thesis Database.

Çiçek Sağlam, A., & Aydoğmuş, M. (2015). Gelişmiş ve gelişmekte olan ülkelerin eğitim sistemlerinin denetim yapıları karşılaştırıldığında Türkiye eğitim sisteminin denetimi ne durumdadır? *Sosyal Bilimler Dergisi*, 2015(25). <https://doi.org/10.12780/uusbd.50788>

Daghe, B. (2018). *A Comparison of Indiana administrator and teacher perceptions of teacher evaluation models* (Publication No. 10830238) [Doctoral dissertation,

Ball State University Muncie, Indiana]. ProQuest Dissertations and Theses Global.

Daley, G., & Kim, L. (2010). A teacher evaluation system that works. *National Institute for Excellence in Teaching*.

Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Association for Supervision & Curriculum Development.

Danielson, C. (2007). *Enhancing professional practice* (2nd ed.). Association for Supervision and Curriculum Development.

Danielson, C. (2011). Evaluations that help teachers learn. *Educational Leadership*, 68(4), 35-39.

Danielson, C. (2013). *The framework for teaching evaluation instruments*. Princeton: The Danielson Group.

Darling-Hammond, L. (2000). Teacher quality and student achievement: a review of state policy evidence. *Education Policy Analysis Archives*, 8(1), 1-44.

Darling-Hammond, L. (2006). *Powerful teacher education: Lessons from exemplary programs*. Jossey-Bass

Darling-Hammond, L. (2010). *Evaluating teacher effectiveness how teacher performance assessments can measure and improve teaching*. Center of American Progress.

Darling-Hammond, L. (2012). *Creating a comprehensive system for evaluating and supporting effective teaching*. Stanford Center for Opportunity Policy in Education.

Darling-Hammond, L. (2014). One Piece of the Whole: Teacher Evaluation as Part of a Comprehensive System for Teaching and Learning. *American Educator*, 38(1), 4-13.

- Darling-Hammond, L., Amrein-Beardsley, A., Haertel, E., & Rothstein, J. (2012). Evaluating Teacher Evaluation. *Phi Delta Kappan*, 93(6), 8-15.
- Darling-Hammond, L., Hyler, M., & Gardner, a. (2017). *Effective Teacher Professional Development*. Learning Policy Institute.
- Dee, T., & Dizon-Ross, E. (2019). School performance, accountability, and waiver reforms: Evidence from Louisiana. *Educational Evaluation and Policy Analysis*, 41(3), 316-349. doi: 10.3102/0162373719849944
- Dee, T., James, J., & Wyckoff, J. (2021). Is effective teacher evaluation sustainable? Evidence from district of Columbia public schools. *Education Finance and Policy*, 16(2), 313-346. https://doi.org/10.1162/edfp_a_00303
- Deiro, J. (2005). *Teachers do make a difference: The Teacher's Guide to Connecting with Students* (1st ed.). Thousand Oaks, CA: Corwin Press.
- Derrington, M. (2011). Changes in teacher evaluation: implications for the principal's work. *Delta Kappa Gamma Bulletin*, 77(3), 51-54.
- Derrington, M., & Brandon, J. (2019). *Differentiated teacher evaluation and professional learning: Policies and practices for promoting career growth*. Palgrave Macmillan.
- Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199. <https://doi.org/10.3102/0013189x08331140>
- De Voto, C., Olson, J., & Gottlieb, J. (2020). Examining diverse perspectives of edTPA policy implementation across states: the good, the bad, and the ugly. *Journal of Teacher Education*, 72(1), 42-55. <https://doi.org/10.1177/0022487120909390>
- Dewey, J. (1910). *How we think?*. Lexington, D.C. Heath and Company.

DiGrazia, D. (2018). *The perceptions of effective teacher evaluation by k-8 teachers and administrators* (Publication No. 10788461) [Doctoral dissertation, Azusa Pacific University]. ProQuest Dissertations and Theses Global.

Directorate General for Teacher Training and Development. (2017). *Teacher Strategy Paper 2017-2023*.

Doherty, K. M. & Jacobs, S. (2013). *State of the States 2013: Connect the Dots--Using Evaluations of Teacher Effectiveness to Inform Policy and Practice*. National Council on Teacher Quality.

Donahue, B. (2016). *The implementation of a new teacher evaluation model: a qualitative case study of how teachers make sense of the Marzano teacher evaluation model* (Publication No. 10090271) [Doctoral dissertation, Rowan University]. ProQuest Dissertations and Theses Global.

Donahue, E., & Vogel, L. (2018). Teacher perceptions of the impact of an evaluation system on classroom instructional practices. *Journal Of School Leadership*, 28(1), 31-55. <https://doi.org/10.1177/105268461802800102>

Donaldson, M. L., Cobb, C. D., LeChasseur, K., Gabriel, R., Gonzalez, R., Woulfin, S., & Makuch, A. (2014). *An evaluation of the pilot implementation of Connecticut's system for educator evaluation and development*. University of Connecticut Center for Education Policy Analysis.

Donaldson, M., & Papay, J. (2015). Teacher evaluation for accountability and development. In H. F. Ladd & M. E. Goertz (Eds.), *Handbook of Research in Education Finance and Policy*. Routledge.

Donaldson, M., & Woulfin, S. (2018). From Tinkering to Going “Rogue”: How Principals Use Agency When Enacting New Teacher Evaluation Systems. *Educational Evaluation and Policy Analysis*, 40(4), 531-556. <https://doi.org/10.3102/0162373718784205>

Education Research and Development Department. (1995). *Öğretmen değerlendirme*. Ankara.

Egan, K. (2010). *Learning in depth: a simple innovation that can transform schooling*. The University of Chicago Press.

Ehren, M., & Baxter, J. (2021). *Trust, accountability, and capacity in education system reform: Global Perspectives in Comparative Education* (pp. 1-23). Routledge.

Evans, L. (2019). *Principals as sense makers: how school leaders make sense of and conduct teacher evaluations in their schools. a phenomenological study of charter school leaders' experiences analyzing, interpreting, and integrating teacher evaluations to improve teacher performance* [Doctoral dissertation, Michigan University]. CMU Graduate Research.

Every Student Succeeds Act. (2019). <http://www.ed.gov/ESSA>

Everson, K. (2016). Value-added modeling and educational accountability. *Review of Educational Research*, 87(1), 35-70. doi: 10.3102/0034654316637199

Firestone, W., & Donaldson, M. (2019). Teacher evaluation as data use: What recent research suggests. *Educational Assessment, Evaluation and Accountability*, 31(3), 289-314. <https://doi.org/10.1007/s11092-019-09300-z>

Firestone, W. A., Nordin, T. L., Shcherbakov, A., Kirova, D., & Blitz, C. L. (2014). *Pilot teacher evaluation program: Year 2 final report*. New Brunswick, NJ: Rutgers Graduate School of Education.

Ford, T., & Hewitt, K. (2020). Better integrating summative and formative goals in the design of next generation teacher evaluation systems. *Education Policy Analysis Archives*, 28, 63. <https://doi.org/10.14507/epaa.28.5024>

Ford, T., Van Sickle, M., Clark, L., Fazio-Brunson, M., & Schween, D. (2016). Teacher self-efficacy, professional commitment, and high-stakes teacher evaluation policy in Louisiana. *Educational Policy*, 31(2), 202-248. <https://doi.org/10.1177/0895904815586855>

Fowler, S. (2001). *Franklin regional growth-oriented teacher evaluation plan opportunities and obstacles an analysis of teachers' perceptions of their first year's experiences in a growth-oriented teacher evaluation plan* (Publication

No. 3041201) [Doctoral dissertation, Duquesne University, Pittsburgh Pennsylvania]. ProQuest Dissertations and Theses Global.

Fullan, M. (2011). *Choosing the wrong drivers for whole system reform*. Centre for Strategic Education.

Fullan, M. & Hargreaves, A. (2016). *Bringing the profession back in: Call to action*. Learning Forward.

Fuller, Thomas C. (2022). *Teachers' perceptions of teacher evaluations at small schools: a qualitative case study* (Publication No. 29214525). [Doctoral dissertation, American College of Education]. ProQuest Dissertations and Theses Global.

Fulton, D. (2019). *Teacher perceptions of the Virginia evaluation process: a program evaluation of Northshore high school's teacher evaluation process* (Publication No. 27546182.) [Doctoral dissertation, The College of William, and Mary in Virginia] ProQuest Dissertations and Theses Global.

Gage, N. (1963). *Handbook of research on teaching*. Rand McNally.

Garet, M., Wayne, A., Brown, S., Rickles, J., Song, M., & Manzeske, D. (2017). *The Impact of Providing Performance Feedback to Teachers and Principals*. Institute of Education Sciences National Center for Education Evaluation and Regional Assistance.

Garrett, R., & Steinberg, M. (2015). examining teacher effectiveness using classroom observation scores. *Educational Evaluation and Policy Analysis*, 37(2), 224-242. <https://doi.org/10.3102/0162373714537551>

Gates Foundation. (2018, June 23). Measures of Effective Teaching (MET) Project - K-12 Education. <https://usprogram.gatesfoundation.org/news-and-insights/articles/measures-of-effective-teaching-project>

Geiger, T., & Amrein-Beardsley, A. (2019). Student perception surveys for K-12 teacher evaluation in the United States: A survey of surveys. *Cogent Education*, 6(1), 1602943. <https://doi.org/10.1080/2331186x.2019.1602943>

Geng, G., Smith, P., Black, P., Budd, Y., & Disney, L. (2019). *Reflective Practice in Teaching: Pre-service Teachers and the Lens of Life Experience* (1st ed.). Springer.

Glazerman, S., Dan, G., Susanna, L., Stephen, R., Douglas, O. S., Grover, J., Whitehurst, & Michelle, C. (2011). *Passing muster: Evaluating teacher evaluation systems*. Brown Center on Education Policy at Brookings.

Gordon, R., Kane, T., & Staiger, D. (2006). *Identifying effective teachers using performance on the job*. Brookings Institution.

Gordon, S., & McGhee, M. (2019). The power of formative evaluation of teaching. In M. Derrington & J. Brandon (Eds.), *Differentiated teacher evaluation and professional learning: policies and practices for promoting career growth*. Palgrave Macmillan.

Göksoy, S. (2014). Hizmet-içi eğitim faaliyetlerinin süreç ve sonuçlarının niteliğine yönelik öğretmen görüşleri. *International Journal of Human Sciences / Uluslararası İnsan Bilimleri Dergisi*, 11(1), 387-402. <https://doi.org/10.14687/ijhs.v11i1.2645>

Graham, L., Berman, J., & Bellert, A. (2015). *Sustainable learning*. Cambridge University Press.

Graham., L., & Berman., J. (2018). *learning intervention: educational casework and responsive teaching for sustainable learning*. Routledge.

Grissom, J., Loeb, S., & Master, B. (2013). Effective instructional time use for school leaders. *Educational Researcher*, 42(8), 433-444. doi: 10.3102/0013189x13510020

Grissom, J., & Youngs, P. (2015). *Improving teacher evaluation systems: Making the most of multiple measures teachers*. College Press.

Guenther, A. (2021). "It should be helping me improve, not telling me I'm a bad teacher": The influence of accountability-focused evaluations on teachers' professional identities. *Teaching and Teacher Education*, 108, 103511. <https://doi.org/10.1016/j.tate.2021.103511>

- Guskey, T. (2019). Grades versus comments: Research on student feedback. *Phi Delta Kappan*, 101(3), 42-47. <https://doi.org/10.1177/0031721719885920>
- Guskey, T., & Link, L. (2022). What teachers really want when it comes to feedback. *Educational Leadership*, 79(7), 42-48.
- Haefele, D. (1993). Evaluating teachers: A call for change. *Journal of Personnel Evaluation in Education*, 7(1), 21-31. <https://doi.org/10.1007/bf00972346>
- Hallinger, P., Heck, R., & Murphy, J. (2014). Teacher evaluation and school improvement: An analysis of the evidence. *Educational Assessment, Evaluation and Accountability*, 26(1), 5-28. <https://doi.org/10.1007/s11092-013-9179-5>
- Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives*, 8, 41. <https://doi.org/10.14507/epaa.v8n41.2000>
- Hanushek, E. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100(1), 84-117. <https://doi.org/10.1086/261808>
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review Of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>
- Hennink, M. (2014). *Focus group discussions*. Oxford University Press.
- Herman, J. L., Heritage, M., & Goldschmidt, P. (2011). *Developing and selecting assessments of student growth for use in teacher evaluation systems*. University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Hernandez, F., & Endo, R. (2017). *Developing and supporting critically reflective teachers*. Sense Publisher.

- Hollweck, T., Curry, A., Smith, K., Dubeau, M. & Kharyati, T. (2019). Prizes and imperfections: Examining teacher evaluation within an induction program in western Québec. In M. Derrington & J. Brandon (Eds.), *Differentiated Teacher Evaluation and Professional Learning: Policies and Practices for Promoting Career Growth*. Palgrave Macmillan.
- Holtzapple, E. (2003). Criterion-related validity evidence for a standards-based teacher evaluation system. *Journal of Personnel Evaluation in Education*, 17(3), 207-219. <https://doi.org/10.1007/s11092-005-2980-z>
- Hopkins, E., Hendry, H., Garrod, F., McClare, S., Pettit, D., & Smith, L. et al. (2016). Teachers' views of the impact of school evaluation and external inspection processes. *Improving Schools*, 19(1), 52-61. <https://doi.org/10.1177/1365480215627894>
- Howard, B. & Gullickson, A. (2009). *Personnel evaluation standards: How to assess systems for evaluating educators (The joint committee on standards for educational evaluation)*. Corwin Press.
- Icel, M. (2008). *Stakeholders' perception on teacher evaluation in the CS teacher evaluation model: exploring teacher and school leader perceptions and a model for improving teacher and student growth* (Publication No. 10979745) [Doctoral dissertation, The Graduate School at the University of Missouri-St. Louis]. ProQuest Dissertations and Theses Global.
- Idapalapati, S. (2019). Towards the fairest teacher evaluation model. *World Journal of Educational Research*, 6(2), 168. <https://doi.org/10.22158/wjer.v6n2p168>
- İlgaz, A. (2011). *Öğretmen performansına denetimin ve yöneticilerin etkileri* (Publication No. 296623) [Master's thesis, Maltepe University]. YÖK Thesis Database.
- Isore, M. (2009). Teacher evaluation: current practices in OECD countries and a literature review: *OECD Education Working Paper No.23*, OECD Publishing.
- Jaffurs, A. (2017). *Teacher perceptions of teacher evaluation using the teacher performance assessment system and factors that contribute to teacher quality, professional growth, and instructional improvement over time* (Publication No.

10640553) [Doctoral dissertation, Graduate School of the University of Maryland-College Park]. ProQuest Dissertations and Theses Global.

Jerald, C. (2009). *Aligned by design how teacher compensation reform can support and reinforce other educational reforms*. Center for American Progress.

Jewell, J. (2017). From inspection, supervision, and observation to value-added evaluation: brief history of U.S. teacher performance evaluations. *Drake Law Review*, 65(2), 363-420.

Jiang, J., Spote, S., & Luppescu, S. (2015). Teacher perspectives on evaluation reform Chicago's REACH students. *Educational Researcher*, 44(2), 105-116.

Johnson, S., & Kardos, S. (2002). Keeping new teachers in mind. *Educational Leadership*, 59(6), 12-16.

Jones, J., Jenkin, M., & Lord, S. (2006). *Developing effective teacher performance*. London: Paul Chapman.

Kahya, O., & Hoşgörür, V. (2020). Çoklu veri kaynağına dayalı performans değerlendirme sistemi hakkında öğretmen görüşleri. *Türkiye Sosyal Araştırmalar Dergisi*, 24(3), 671-681.

Kane, T., & Staiger, D. (2012). *Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains*. Seattle: Bill and Melinda Gates Foundation.

Kane, T., Kerr, K., & Pianta, R. (2014). *Designing teacher evaluation systems* (1st ed.). San Francisco: Josey-Bass.

Kaplan, İ. (2019). *Öğretmenlerin sürekli mesleki gelişimini destekleyen denetim modeli* (Publication No. 559019) [Doctoral dissertation, Gazi University]. YÖK Thesis Database.

Karasolak, K., Tanrıseven, I., & Yavuz-Konakman, G. (2012). Öğretmenlerin hizmetiçi eğitim etkinliklerine ilişkin tutumlarının belirlenmesi. *Kastamonu Eğitim Dergisi*, 21(3), 997-1010.

Kennedy, M. (2005). *Inside teaching*. London. Harvard University Press.

Kennedy, M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945-980. <https://doi.org/10.3102/0034654315626800>

Kim, J., & Sun, M. (2021). The implementation and potential effects of teacher evaluation under local control. *School Effectiveness and School Improvement*, 32(2), 279-305. <https://doi.org/10.1080/09243453.2020.1841657>

Kimball, S. (2001). *Innovations in teacher evaluation: case studies of two school districts with teacher evaluation systems based on the framework for teaching* (Publication No. 3033269) [Doctoral dissertation, University of Wisconsin-Madison]. ProQuest Dissertations and Theses Global.

Kimball, S., & Milanowski, A. (2009). Examining Teacher evaluation validity and leadership decision making within a standards-based evaluation system. *Educational Administration Quarterly*, 45(1), 34-70. <https://doi.org/10.1177/0013161x08327549>

Kini, T., & Podolsky, A. (2016). *Does teaching experience increase teacher effectiveness? A Review of the Research*. Palo Alto: Learning Policy Institute.

Kitamura, Y., Omomo, T., & Katsuno, M. (2019). *Education in Japan: A comprehensive analysis of education reforms and practices*. Springer.

Klein, A. (2020, September 12). Every student succeeds act: An ESSA overview. *Education Week*. <https://www.edweek.org/policy-politics/the-every-student-succeeds-act-an-essa-overview/2016/03>

Kohn, A. (2000). *What to look for in a classroom*. San Francisco: Jossey-Bass-Publication.

Kraft, M.A. & Gilmour, A.F. (2017). Revisiting the widget effect: Teacher evaluation reforms and the distribution of teacher effectiveness. *Educational Researcher*, 46(5), 234-249.

- Kraft, M., & Papay, J. (2014). Can professional environments in schools promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*, 36(4), 476-500. <https://doi.org/10.3102/0162373713519496>
- Kratz, H. (1896). Characteristics of the Best Teacher as Recognized by Children. *The Pedagogical Seminary*, 3(3), 413-460. <https://doi.org/10.1080/08919402.1896.9943604>
- Kritt, D. (2018). *Constructivist education in an age of accountability*. Palgrave Macmillan.
- Krueger, R., & Casey, M. (2015). *Focus groups: A Practical guide for applied research* (5th ed.). Sage.
- Kurum, G., & Cinkir, S. (2019). An authentic look at evaluation in education: a school self-evaluation model supporting school development. *Eurasian Journal of Educational Research*, 19(83), 1-34. <https://doi.org/10.14689/ejer.2019.83.12>
- La Masa, M. (2005). *An elementary school's perceptions of the effectiveness of teacher evaluation to enhance teacher practice* (Publication No. 3180461) [Doctoral dissertation, University of Southern California]. ProQuest Dissertations and Theses Global
- Lavigne, A. (2014). Exploring the Intended and unintended consequences of high-stakes teacher evaluation on schools, teachers, and students. *Teachers' college record: The Voice of Scholarship in Education*, 116(1), 1-29. <https://doi.org/10.1177/016146811411600103>
- Lavigne, A., & Good, T. (2019). *Enhancing teacher education, development, and evaluation: Lessons learned from educational reform*. Routledge.
- LeCompte, M. (2009). Trends in research on teaching: An historical and critical overview. In L. Saha & A. Dworkin (Eds.), *International Handbook of Research on Teachers, and Teaching* (pp. 25-61). Springer.
- Lee, L., & Cousins, J. (1995). Participation in Evaluation of Funded School Improvement: Effects and Supporting Conditions. In J. Cousins & L. Earl

(Eds.), *Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning*. Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning.

Leedy, P., & Ormrod, J. (2016). *Practical research: Planning and design* (11th ed.). Pearson.

Leighton, J. (2020). Cognitive diagnosis is not enough the challenge of measuring learning with classroom assessments. In S. Brookhart & J. McMillan (Eds.), *Classroom Assessment and Educational Measurement* (1st ed., pp. 27-40). New York: Routledge.

Lewis, S., & Hogan, A. (2016). Reform first and ask questions later? The implications of (fast) schooling policy and ‘silver bullet’ solutions. *Critical Studies in Education*, 60(1), 1-18. <https://doi.org/10.1080/17508487.2016.1219961>

Lieberman, A. (2000). Networks as learning communities. *Journal Of Teacher Education*, 51(3), 221-227. <https://doi.org/10.1177/0022487100051003010>

Lieberman, A., & Miller, L. (2014). Teachers as professionals: Evolving definitions of staff development. In L. E. Martin, S. Kragler, D. J. Quatroche, & K. L. Bauserman (Eds.), *Handbook of professional development in education: Successful models and practices, PK-12* (pp. 3–21). Guilford Press.

Liebowitz, D., & Porter, L. (2019). The effect of principal behaviors on student, teacher, and school outcomes: a systematic review and meta-analysis of the empirical literature. *Review of Educational Research*, 89(5), 785-827. doi: 10.3102/0034654319866133

Lillejord, S., Elstad, E., & Kavli, H. (2018). Teacher evaluation as a wicked policy problem. *Assessment In Education: Principles, Policy & Practice*, 25(3), 291-309. <https://doi.org/10.1080/0969594x.2018.1429388>

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Sage Publication.

Little, J. (2003). Inside teacher community: Representations of classroom practice. *Teachers College Record: The Voice of Scholarship in Education*, 105(6), 913-945. <https://doi.org/10.1111/1467-9620.00273>

- Liu, L. & Zhang, Y. (2014). *Enhancing teachers' professional development through reflective teaching*. *TPLS*, 4(11).
- Liu, S., & Zhao, D. (2013). Teacher evaluation in China: latest trends and future directions. *Educational Assessment, Evaluation and Accountability*, 25(3), 231-250. <https://doi.org/10.1007/s11092-013-9168-8>
- Looney, J. (2011). Developing high-quality teachers: Teacher evaluation for improvement. *European Journal of Education*, 46(4), 440-455. <https://doi.org/10.1111/j.1465-3435.2011.01492.x>
- Lynda, T., Matt, M., & Ken, S. (2021). Making classroom observations matter. *Educational Leadership*, 78(7), 56-62.
- Marshall, K. (2013). *Rethinking teacher supervision and evaluation*. San Francisco: Jossey-Bass.
- Marshall, J., & Horton, R. (2011). The relationship of teacher-facilitated, inquiry-based instruction to student higher-order thinking. *School Science and Mathematics*, 111(3), 93-101. doi: 10.1111/j.1949-8594.2010.00066.x
- Marshall, C., & Rossman, G. (2015). *Designing qualitative research* (6th ed.). Sage.
- Marx, R. (2014). Reforming again: Now teachers. *Teachers College Record: The Voice of Scholarship in Education*, 116(1), 1-15. <https://doi.org/10.1177/016146811411600107>
- Marzano, R., Frontier, T., & Livingston, D. (2011). *Effective supervision: supporting the art and science of teaching*. ASCD.
- Marzano, R. (2012). The two purposes of teacher evaluation. *Educational Leadership*, 70(3), 14-19.
- Marzano, R., & Simms, J. (2014). *Questioning sequences in the classroom*. Marzano Research Laboratory.
- Marzano, R., & Toth, M. (2013). *Teacher evaluation that makes a difference*. ASCD.

Mason, J. (2002). *Qualitative researching* (2nd ed.). Sage.

Master, B. (2014). Staffing for success. *Educational Evaluation and Policy Analysis*, 36(2), 207-227. <https://doi.org/10.3102/0162373713506552>

Maxwell, J. (2013). *Qualitative research design: An interactive approach* (3rd ed.). SAGE.

McAlister, A., Lee, D., Ehlert, K., Kajfez, R., Faber, C., & Kennedy, M. (2017). Qualitative coding: An approach to assess inter-rater reliability. *2017 ASEE Annual Conference & Exposition Proceedings*. <https://doi.org/10.18260/1-2--28777>

McClure, A. (2008). *Teacher evaluation: Pre-service to in-service* [Doctoral dissertation, University of Toronto].

McCombs, B. (1997). Self-assessment and reflection: Tools for promoting teacher changes toward learner-centered practices. *NASSP Bulletin*, 81(587), 1-14. <https://doi.org/10.1177/019263659708158702>

McGuinn, P. (2012). *The State of teacher evaluation reform: State education agency capacity and the implementation of new teacher-evaluation systems*. Center of American Progress.

McQueen, D. (2022). *A qualitative case study on the perceptions of teachers' current teacher evaluation system at a small rural public school district in Northern California* (Publication No. 29063667) [Doctoral dissertation, Northcentral University]. ProQuest Dissertations and Theses Global.

MEBBİS. (2018, September 19). *MEB Öğretmen Performansı Değerlendirme*. <https://mebbisyd.meb.gov.tr/>.

Memduhoğlu H. B. (2012). Öğretmen, yönetici, denetmen ve öğretim üyelerinin görüşlerine göre Türkiye’de eğitim denetimi sorunsalı. *Kuram ve Uygulamada Eğitim Yönetimi*, 12(1), 135-156.

- Medley, D. (1979). The effectiveness of teachers. In P. Peterson & H. Walberg (Eds.), *Research on Teaching: Concepts, Findings, and Implications (Series on contemporary educational issues)* (1st ed.).
- Miles, M., & Huberman, A. (1994). *Qualitative data analysis*. Sage.
- Miles, M., Huberman, A., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.
- Millman, J., & Darling-Hammond, L. (1991). *The new handbook of teacher evaluation*. Corwin Press.
- Ministry of National Education. (2014). *Milli Eğitim Bakanlığı Rehberlik ve Denetim Başkanlığı ile Maarif Müfettişleri Başkanlıkları Yönetmeliği*. Official newspaper.
- Ministry of National Education. (2017). *Öğretmenlik mesleği genel yeterlikleri*. Ankara: Öğretmen Yetiştirme ve Geliştirme Genel Müdürlüğü.
- Ministry of National Education. (2018). *Öğretmen Performans Değerlendirme Yönetmelik Taslağını*.
- Moss, J. (2015). *The Danielson model of teacher evaluation: exploring teacher perceptions concerning its value in shaping and improving instructional practice* (Publication No. 3717968) [Doctoral dissertation, Seton Hall University]. ProQuest Dissertations and Theses Global.
- Murphy, J., Hallinger, P., & Heck, R. (2013). Leading via teacher evaluation. *Educational Researcher*, 42(6), 349-354. <https://doi.org/10.3102/0013189X13499625>
- Muskin, J. (2015). From good ideas to good practice: putting teachers at the center of education improvement, where they belong. *Educational Assessment, Evaluation and Accountability*, 27(1), 93-102. <https://doi.org/10.1007/s11092-015-9216-7>
- Murtagh, L. (2014). The motivational paradox of feedback: teacher and student perceptions. *The Curriculum Journal*, 25(4), 516-541. <https://doi.org/10.1080/09585176.2014.944197>

National Council on Teacher Quality. (2011). *State of the states: trends and early lessons on teacher evaluation and effectiveness policies*. http://www.nctq.org/p/publications/docs/nctq_stateOfTheStates.pdf

National Education Association. (2012). *Teacher assessment and evaluation: The national education association's framework for transforming education systems to support effective teaching and improve student learning*. http://www.nea.org/assets/docs/HE/TeachrAssmntWhtPaperTransform10_2.pdf

National Institute for Excellence in Teaching, (2012). *Career teacher handbook: Tap Instructional rubrics*. TAP™: The System for Teacher and Student Advancement.

National Institute for Excellence in Teaching, (2022, August 6). *The TAP system for teacher and student advancement*. <https://www.niet.org/our-work/our-services/show/the-tap-system-for-teacher-and-student-advancement>

Nelson, J. (2015). *Teacher evaluation: teachers' past experiences and their responses to impending reform* (Publication No. 3704781) [Doctoral dissertation, Illinois State University]. ProQuest Dissertations and Theses Global.

Netolicky, D. (2020). *Transformational professional learning: making a difference in schools*. Routledge.

Nichols, S., & Berliner, D. (2008). *Collateral damage*. Harvard Education Press.

Nikolic, V., & Cabaj, H. (2003). *Am I teaching well? self-evaluation strategies for effective teachers*. Pippin Publishing Corporation.

Nilsen, P. (2006). *The teacher evaluation process: impacts and implications on teacher behavior in the classroom a case study on the link between teacher evaluation and the student success in a California elementary school* (Publication No. 3233804) [Doctoral dissertation, University of Southern California]. ProQuest Dissertations and Theses Global.

Organization for Economic Co-operation and Development. (2009a), *Creating effective teaching and learning environments: First results from TALIS*. Paris: OECD Publishing.

Organization for Economic Co-operation and Development. (2009b). *Teacher evaluation a conceptual framework and examples of country practices*. Paris: OECD Publishing.

Organization for Economic Co-operation and Development. (2013a). *Teachers for the 21st century using evaluation to improve teaching*. OECD Publishing.

Organization for Economic Co-operation and Development. (2013b). *Synergies for better learning: An international perspective on evaluation and assessment*. OECD Publishing.

Organization for Economic Co-operation and Development. (2014). *A teachers' guide to TALIS 2013: Teaching and learning international survey*. Paris: OECD Publishing.

Organization for Economic Co-operation and Development. (2018), *Teaching for the Future: Effective classroom practices to transform*. Paris: OECD Publishing.

Organization for Economic Co-operation and Development. (2019). *OECD Reviews of Evaluation and Assessment in Education: Student Assessment in Turkey: OECD Reviews of Evaluation and Assessment in Education*. Paris: OECD Publishing.

Organization for Economic Co-operation and Development. (2020a). *Education Policy Outlook: Turkey*. Paris: OECD Publishing.

Organization for Economic Co-operation and Development. (2020b). *TALIS 2018 results: Teachers and school leaders as lifelong learners*. Paris: OECD Publishing.

Office for Standards in Education (Ofsted). (2018). *Six models of lesson observation: an international perspective*. Manchester.

- Ornstein, A. (1985). Research on teaching: Issues and trends. *Journal Of Teacher Education*, 36(6), 27-31. <https://doi.org/10.1177/002248718503600604>
- Ortlipp, M. (2015). Keeping and using reflective journals in the qualitative research process. *The Qualitative Report*. 13(4), 695-705 <https://doi.org/10.46743/2160-3715/2008.1579>
- Özbek, Ö., & Taneri, P. (2019). Teacher performance appraisal in Turkey: An alternate model. *European Journal of Education*, 5(15), 222-228.
- Özkan, H., & Çelikten, Y. (2017). Milli eğitim sisteminin örgütsel yapısı ve maarif müfettişleri alt sisteminin işleyişi. *OPUS Uluslararası Toplum Araştırmaları Dergisi*, 7(13), 965-990. <https://doi.org/10.26466/opus.321994>
- Papay, J. (2012). Refocusing the debate: assessing the purposes and tools of teacher evaluation. *Harvard Educational Review*, 82(1), 123-141. <https://doi.org/10.17763/haer.82.1.v40p0833345w6384>
- Passe, A. (2015). *Evaluating and supporting early childhood teachers* (1st ed.). Redleaf Press.
- Patton, M. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE.
- Paufler, N., King, K., & Zhu, P. (2020). Promoting professional growth in new teacher evaluation systems: Practitioners' lived experiences in changing policy contexts. *Studies In Educational Evaluation*, 65, 100873. <https://doi.org/10.1016/j.stueduc.2020.100873>
- Pehlivan, İ., Demirbaş, A., & Eroğlu, E. (2001). *Öğretmen performansı değerlendirme ve sicil raporları: Education research and development department research report*. Ankara: Milli Eğitim Basımevi.
- Pianta, R. (2012). *Implementing observation protocols lessons for K-12 education from the field of early childhood*. Center for American Progress.

- Pianta, R. & Hambre, B. (2015). Implementing rigorous observation of teachers: synchronizing theory with systems for implementation and support. In J. Grissom & P. Youngs (Eds.), *Improving Teacher Evaluation Systems: Making the Most of Multiple Measures* (pp. 22-30). Teachers College Press.
- Popham, W. (2013). *Evaluating America's teachers*. SAGE Publications.
- Puchta, C., & Potter, J. (2004). *Focus Group Practice*. Sage.
- Putman, H., Walsh, K., & Ross, E. (2018). *Making a difference: Six places where teacher evaluation systems are getting results*. National Council on Teacher Quality.
- Ravitch, S., & Mittenfelner, N. (2016). *Qualitative research: Bridging the Conceptual, Theoretical, and Methodological*. Sage.
- Reddy, L., Dudek, C., Kettler, R., Kurz, A., & Peters, S. (2016). Measuring educators' attitudes and beliefs about evaluation: Construct validity and reliability of the teacher evaluation experience scale. *Educational Assessment*, 21(2), 120-134.
- Reinhorn, S., Johnson, S., & Simon, N. (2017). Investing in development: Six high-performing, high-poverty schools implement the Massachusetts teacher evaluation policy. *Educational Evaluation and Policy Analysis*, 39(3), 383-406. <https://doi.org/10.3102/0162373717690605>
- Roberts, K., Dowell, A., & Nie, J. (2019). Attempting rigour and replicability in thematic analysis of qualitative research data; a case study of codebook development. *BMC Medical Research Methodology*, 19(1). <https://doi.org/10.1186/s12874-019-0707->
- Robertson-Kraft, C., & Zhang, R. (2016). Keeping great teachers: A case study on the impact and implementation of a pilot teacher evaluation system. *Educational Policy*, 32(3), 363-394. <https://doi.org/10.1177/0895904816637685>
- Robinson, K., & Aronica, L. (2016). *Creative schools: The grassroots revolution that's transforming education*. Penguin Books.

- Rodriguez, B., & Garza, S. (2014). Test review: Classroom assessment scoring system manual: Toddler. *Journal of Psychoeducational Assessment*, 32(6), 573-576. <https://doi.org/10.1177/0734282914523085>
- Ross, E., & Walsh, K. (2019). *State of the states 2019: Teacher and principal evaluation policy*. National Council on Teacher Quality
- Rossman, G., & Rallis, S. (2011). *Learning in the field: An introduction to qualitative research* (3rd ed.). Sage.
- Royal, C. & Tossman, M. (2009). *Philadelphia's teacher appraisal system needs improvement*. Philadelphia: Research for Action.
- Rucinski, M. and Diersing, C., (2014). America's teacher evaluation system revolution. *Harvard Political Review*.
- Rumage, J. (2012). *Teacher evaluation and professional growth*. College of Saint Elizabeth.
- Rushton, I. & Suter, M. (2012). *Reflective practice for teaching in lifelong learning*. Maidenhead: McGraw-Hill Education.
- Ryu, J. (2020). Supporting professional capacity of teachers through teacher evaluation: Associations of school climate, teacher evaluation, and professional capacity in four countries (Publication No. 9203) [Doctoral dissertation, The University of Minnesota]. University of Minnesota Digital Conservancy.
- Saldaña, J. (2016). *The coding manual for qualitative researchers*. Sage.
- Sanders, W., Wright, S., & Horn, S. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal Of Personnel Evaluation in Education*, 11(1), 57-67.
- Shenton, A. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, 22(2), 63-75. <https://doi.org/10.3233/efi-2004-22201>

Shinkfield, A., & Stufflebeam, D. (1997). *Teacher evaluation*. Kluwer Academic Publishers.

Shulha, L., Whitmore, E., Cousins, J., Gilbert, N., & Al Hudib, H. (2015). Introducing evidence-based principles to guide collaborative approaches to evaluation. *American Journal of Evaluation*, 37(2), 193-215. <https://doi.org/10.1177/1098214015615230>

Shulman, L. (1986). Those who understand knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14. doi: 10.3102/0013189x015002004

Sinnema, C., & Robinson, V. (2007). The leadership of teaching and learning: implications for teacher evaluation. *Leadership And Policy in Schools*, 6(4), 319-343. <https://doi.org/10.1080/15700760701431603>

Sloat, E., Amrein-Beardsley, A., & Sabo, K. (2017). Examining the factor structure underlying the TAP system for teacher and student advancement. *AERA Open*, 3(4), 233285841773552. <https://doi.org/10.1177/2332858417735526>

Sporte, S., Stevens, W., Healey, K., Jiang, J., & Hart, H. (2013). *Teacher evaluation in practice implementing Chicago's REACH students*. UChicago CCSR's publications.

Stallings, J. (1985). A study of implementation of Madeline Hunter's model and its effects on students. *The Journal of Educational Research*, 78(6), 325-337. <https://doi.org/10.1080/00220671.1985.10885626>

Steinberg, M., & Kraft, M. (2017). The sensitivity of teacher performance ratings to the design of teacher evaluation systems. *Educational Researcher*, 46(7), 378-396. <https://doi.org/10.3102/0013189x17726752>

Steiner, L. (2010). *Using competency-based evaluation to drive teacher excellence: Lessons from Singapore* (pp. 1-18). Chapel Hill: Public Impact.

Stecher, B., Holtzman, D., Garet, M., Hamilton, L., Engberg, J., & Steiner, E. et al. (2018). *Improving teaching effectiveness: final report the intensive partnerships for effective teaching through 2015–2016*. RAND.

- Steinberg, M., & Donaldson, M. (2016). The new educational accountability: understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy*, 11(3), 340-359. https://doi.org/10.1162/edfp_a_00186
- Stronge, J. (2006). *Evaluating teaching: A guide to current thinking and best practice*. Corwin Press.
- Stronge, J. (2018). *Qualities of effective teachers* (3rd ed.). ASCD.
- Stronge, J., & Caine Tonneson, V. (2018). *Teacher effectiveness performance evaluation system handbook*. Stronge & Associates.
- Stronge, J. H., & Tucker, P. D. (2003). *Developing and implementing a teacher evaluation system*. Eye on Education.
- Süzen, A. (2007). *İnsan kaynakları yönetim süreçleri çerçevesinde öğretmen değerlendirmesinde performans değerlendirme: özel bir ilköğretim okulundaki sınıf öğretmenlerinin görüşleri* (Publication No. 210090) [Master's thesis, Eskişehir Anadolu University]. YÖK Thesis Database.
- Şahan, H. (2011). Öğretmenlerin öğretim hizmetinin niteliğini etkileyen değişkenler açısından kendilerini değerlendirmeleri. *Balıkesir University Journal Of Social Sciences Institute*, 14(26), 36-45.
- Taubman, P. (2009). *Teaching by numbers: Deconstructing the discourse of standards and accountability in education*. Routledge.
- Taş, İ. (2020). Teacher evaluation in various countries. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 22(2), 311-325.
- Taş, İ., & Bıkmaz, F. (2019). Bir öğretmen değerlendirme modeli önerisi. *Hacettepe University Journal of Education*, 1-25. <https://doi.org/10.16986/huje.2019053980>

Tarhan, H., Karaman, A., Kemppinen, L., & Aerila, J. (2019). Understanding teacher evaluation in Finland: A professional development framework. *Australian Journal of Teacher Education*, 44(4).

Taylor, E.S, & Tyler, J.H. (2012). Can teacher evaluation improve teaching? *Education Next*, 12(4).

TEDMEM. (2018). *Öğretmen Performans Değerlendirme ve Aday Öğretmenlik İş ve İşlemleri Yönetmelik Taslağı Üzerine Değerlendirmeler*.

The New Teacher Project, (2010). *Teacher evaluation 2.0*. TNTP Publications.

The New Teacher Project, (2016). *Accountability under ESSA: how states can design systems that advance equity and opportunity*. TNTP Publications.

Toch, T., & Rothman, R. (2008). *Rush to Judgment*. Washington: Education Sector.

Tracy, S. (2020). *Qualitative research methods: collecting evidence, crafting analysis, communicating impact* (2nd ed.). Wiley-Blackwell.

Tucker, P., & Stronge, J. (2005). *Linking teacher evaluation and student learning*. Association for Supervision and Curriculum Development.

Tuytens, M., & Devos, G. (2013). How to activate teachers through teacher evaluation? *School effectiveness and school improvement*, 25(4), 509-530. <https://doi.org/10.1080/09243453.2013.842601>

Türkoğlu, M. (2015). *Öğretmen hesap verebilirliği: özel bir okulda durum çalışması* (Publication No. 391157) Doctoral dissertation, Eskişehir Osmangazi University].

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2017). *Accountability in education: Meeting our Commitments* .

U.S. Department of Education (2007). *State and Local Implementation of the No Child Left Behind Act, Volume II—Teacher Quality Under NCLB*. file:///C:/Users/Emel/Downloads/RAND_RP1283.pdf

- Vandermolen, R. M., & Mey-Looze, C. L. (2021). A regional case study: Effective administrator feedback to inform teacher performance. *Delta Kappa Gamma Bulletin*, 87(5), 38–49
- Viennet, R. and B. Pont (2017), *Education policy implementation: A literature review and proposed framework*, OECD Education Working Papers, No. 162, OECD Publishing, Paris, <https://doi.org/10.1787/fc467a64-en>.
- Von der Embse, N., Schoemann, A., Kilgus, S., Wicoff, M., & Bowler, M. (2016). The influence of test-based accountability policies on teacher stress and instructional practices: a moderated mediation model. *Educational Psychology*, 37(3), 312-331. <https://doi.org/10.1080/01443410.2016.1183766>
- Walker, T. T. (2014). *Teacher evaluation: Key components the teacher evaluation: key components that positively impact teacher effectiveness and student achievement* (Publication No. 3628338) [Doctoral dissertation, University of Southern California]. ProQuest Dissertations and Theses Global.
- Walsh, J., & Sattes, B. (2015). *Questioning for classroom discussion: Purposeful speaking, engaged listening, deep thinking*. ASCD.
- Walsh, K., Joseph, N., Lakis, K., & Lubell, S. (2017). *Running in place: How new teacher evaluations fail to live up to promises*. National Council on Teacher Quality. Retrieved from http://www.nctq.org/dmsView/Final_Evaluation_Paper
- Washington Office of Superintendent of Public Instruction (2022, May 4). *CEL 5D+ Teacher evaluation rubric 3.0* <https://www.k12.wa.us/educator-support/teacherprincipal-evaluation-program/frameworks-and-rubrics/cel-5d-teacher-evaluation-rubric-30>.
- Webb, R. (2006). *Changing teaching and learning in the primary school*. Maidenhead: Open University Press.
- Wei, R. C., Darling-Hammond, L., Andree, A., Richardson, N., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. Dallas, TX: National Staff Development Council.

Weisberg, D., Sexton, S., Muhern, J., & Keeling, D. (2009). *The widget effect*. The New Teacher Project.

Williams, J., & Engel, L. (2013). How do other countries evaluate teachers? *Phi Delta Kappan*, 94(4), 53-57. <https://doi.org/10.1177/003172171209400414>

Winslow, R. (2015). *Administrative feedback following classroom observations as part of a Danielson-based teacher evaluation system: teacher and administrator perceptions* (Publication No. 3705399). [Doctoral dissertation, Graduate School Southern Illinois University Edwardsville]. ProQuest Dissertations and Theses Global.

Wind, S., & Jones, E. (2019). Not just generalizability: A case for multifaceted latent trait models in teacher observation systems. *Educational Researcher*, 48(8), 521-533. doi: 10.3102/0013189x19874084

White, K. (2016). *Teacher communication: A guide to relational, organizational, and classroom communication*. Rowman & Littlefield.

William, D. (2018). *Creating the Schools Our Children Need*. Learning Sciences International.

Winslow, R. (2015). *Administrative feedback following classroom observations as part of a Danielson-based teacher evaluation system: teacher and administrator perceptions* (Publication No. 3705399). [Doctoral dissertation, Graduate School Southern Illinois University Edwardsville]. ProQuest Dissertations and Theses Global.

Wise, A., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1984). *Teacher evaluation: A study of effective practices*. RAND.

Wong, M., & Li, H. (2010). From external inspection to self-evaluation: A study of quality assurance in Hong Kong kindergartens. *Early Education & Development*, 21(2), 205-233. <https://doi.org/10.1080/10409281003638725>

Yılmaz, A. (2017). *Eğitim denetimi sistemleri bakımından Türkiye ile PISA'da başarı gösteren bazı ülkelerin karşılaştırılması* (Publication No. 454361) [Master's thesis, Çanakkale Onsekiz Mart University]. YÖK Thesis Database.

Yin, R. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage Publications.

Zarro, J. (2005). *The impact of teacher evaluation on teacher practice: a case study* (Publication No. 3180328) [Doctoral dissertation, University of Southern California]. ProQuest Dissertations and Theses Global.

Zepeda, S., & Lanoue, P. (2017). *Conversation walks: Improving Instructional Leadership*. ASCD.



APPENDICES

A. TEACHER INTERVIEW SCHEDULE

ÖĞRETMEN GÖRÜŞME FORMU

Tarih:

Katılımcı:

Süre:

Sayın ...

Orta Doğu Teknik Üniversitesi Eğitim Bilimleri Bölümü, Eğitim Programları ve Öğretimi programında doktora öğrencisiyim. Doktora tezimi bu okulda görev alan öğretmenlerin, yöneticilerin ve uzmanların, öğretmen değerlendirme hakkındaki görüşlerini alarak ihtiyaçlarını anlamak, önerileri ve katılımları doğrultusunda öğretmenlerin mesleki gelişimlerini destekleyecek bir değerlendirme modeli hazırlamak için yürütmekteyim. Bu bağlamda, siz değerli öğretmenlerin görüşlerini almak benim için çok önemli.

Bu çalışmaya katıldığınız ve zaman ayırdığınız için çok teşekkür ederim. Bu röportaj, kurumunuzda etkili bir öğretmen değerlendirme modeli oluşturma süreci hakkındaki görüşlerinizi paylaşmak için iyi bir fırsat olabilir.

Görüşmenin genel özellikleri ve gizliliği hakkında bazı hatırlatmalar yapmak istiyorum:

- Bu konuşma sadece araştırma amaçlı kullanılacaktır. Adınız hiçbir rapor vb. belgede kullanılmayacak ve araştırmacı tarafından gizli tutulacaktır.*
- Görüşmedeki ayrıntıları ve önemli bir şeyi kaçırmamam için kaydetmek istiyorum. Kayıtlar tamamen gizli tutulacaktır.*
- Konuşmamızın hiçbir bölümünde, sizi yanıltacak veya size zarar verebilecek herhangi bir şey yoktur. Bununla birlikte, görüşmenin herhangi bir aşamasında, istemiyorsanız, mülakat derhal feshedilecek ve tüm kayıt iptal edilecek.*

Görüşme yaklaşık 40 dakika sürecektir. Başlamadan önce sormak istediğiniz bir şey var mı?

Tanımlayıcı Bilgiler

- a. Kaç yıldır öğretmenlik yapıyorsunuz?
- b. Bu okulda kaç yıldır öğretmenlik yapıyorsunuz? Daha önce hangi kurumda çalıştınız?
- c. Okulda idari bir göreviniz var mı? Daha önce idari bir göreviniz oldu mu? Örneğin, Zümre başkanlığı, müdür yardımcılığı vb.
- d. Öğrenim düzeyinizden bahseder misiniz?
- e. Bu okulda ya da daha önce çalıştığınız kurumlarda bir değerlendirme sürecinden geçtiniz mi?

Görüşme Soruları

1. Öğretmen değerlendirme süreci denildiğinde aklınıza neler geliyor?
(Bir öğretmen değerlendirme sürecini nasıl tanımlarsınız?)

2. Okulunuzda yapılan değerlendirme süreçleri ile ilgili örnekler verebilir misiniz?
Yapılan değerlendirmeler ile ilgili bir deneyiminizi paylaşır mısınız?

3. Okulunuzda yürütülen öğretmen değerlendirme süreci ve bu sürecin nasıl olması gerektiği hakkında biraz konuşalım.

A1. Sizi kimler değerlendiriyor? Değerlendirmeyi yapan kişilerin bu alandaki yeterlikleri hakkında neler düşünüyorsunuz?

Sonda: Yöneticiler, veliler, öğrenciler, aynı branştaki öğretmenler, farklı branştaki öğretmenler, kurum dışından kişiler vb.

A2. Bu kişiler tarafından değerlendirilmenin, değerlendirme sürecine olumlu etkileri nelerdir? Olumsuz etkileri nelerdir?

A3. Sizi kimler değerlendirmelidir? Değerlendirmeyi yapan kişiler hangi yeterliklere ve özelliklere sahip olmalıdır?

B1. Değerlendirmeler ne sıklıkta yapılıyor? Ne zaman yapılıyor?

B2. Değerlendirme sıklığının olumlu yönleri nelerdir? Olumsuz yönleri nelerdir?

B3. Değerlendirmeler ne sıklıkta yapılmalıdır? Ne zaman yapılmalı?

C1. Okulunuzdaki değerlendirmede ne tür değerlendirme araçları kullanılıyor?

Sonda: Gözlem, görüşme, öz-değerlendirme, uzmanlar tarafından hazırlanmış standart formlar, anketler vb.

C2. Kullanılan bu araçların, değerlendirme sürecine olumlu katkıları nelerdir? Bu araçları kullanmanın süreçte sebep olduğu olumsuzluklar nelerdir?

C3. Değerlendirme modelinde ne tür değerlendirme araçları kullanılmalıdır?

D1. Okulunuzdaki değerlendirmelerde ne tür veri kaynakları kullanılıyor?

Sonda: Ders planları, derslerde kullandığınız araç ve gereçler, öğrencilerin başarı göstergeleri (karne notu, sınav sonuçları vb.), gözlem sonuçları, görüşme sonuçları, öz-değerlendirme sonuçları vb.

D2. Bu veri kaynaklarının kullanılmasının değerlendirme sürecine olumlu etkileri nelerdir? Olumsuz etkileri nelerdir?

D3. Değerlendirme modelinde ne tür veri kaynakları kullanılmalıdır?

E1. Hangi yeterlikleriniz değerlendiriliyor?

Sonda:

Derse hazırlık yapma (dersi planlama, sınıf yönetimi, kullanılan araç-gereç ve malzemeler, farklılıkları dikkate alma vb.)

Öğretimi gerçekleştirme (etkili öğrenme ortamı oluşturma, farklılıkları dikkate alma, üst düzey düşünme becerilerini destekleme vb.)

Öğretimi izleme (öğrencilerin gelişimlerini takip etme, eksikliklerini belirleme ve telafi etme)

Okul içindeki diğer sorumlulukları yerine getirme (okuldaki diğer görevleriniz, zümre içindeki görevlerini vb.)

İlişkiler (öğretmenlerle, velilerle, yöneticilerle etkili iletişim kurmak, iş birliği içinde çalışmak, vb.)

Mesleki gelişim (eğitime katılma, yayınları takip etme, gelişmelerden haberdar olma vb.)

E2. Değerlendirmenin bu yeterliklere göre yapılmasının olumlu etkileri nelerdir? Değerlendirmenin bu özelliklere göre yapılmasının sebep olduğu olumsuzluklar nelerdir?

E3. Hangi yeterlikleriniz değerlendirilmelidir?

F1. Okulunuzda yürütülen değerlendirmelerde objektiflik (tarafsızlık) nasıl sağlanıyor?

(Değerlendirme sonuçlarına güveniyor musunuz?)

F2. Değerlendirmelerin objektifliği (tarafsızlığı) nasıl sağlanmalıdır?

G1. Değerlendirme sonuçları ne amaçla kullanılıyor?

(Ne amaçla değerlendiriliyorsunuz?)

Sonda: Kuruma ilişkin raporlamalar (faaliyet, memnuniyet, gelişim vb.) yapma, güçlü ya da geliştirilmesi gereken yönleriniz belirlemek, uyarı ya da ceza vermek, öğrenci başarısını arttırmak, profesyonel gelişimizi desteklemek)

G2. Değerlendirme sonuçlarının bu amaçlarla kullanılmasının olumlu yönleri nelerdir? Olumsuzlukları nelerdir?

G3. Değerlendirmeden elde edilen veriler ne amaçla kullanılmalıdır?

(Bu verilere dayalı olarak neler yapılmalıdır ve ne tür kararlar alınmalıdır?)

H1. Değerlendirme sonuçlarından nasıl haberdar ediliyorsunuz? Ne sıklıkla haberdar ediliyorsunuz?

Sonda: Yüz yüze görüşmelerle, yazılı iletişim yoluyla, telefonla, düzenli olarak, rastgele

H2. Bu şekilde haberdar edilmenin olumlu yönleri nelerdir? Olumsuz yönleri nelerdir?

H3. Değerlendirme sonuçları nasıl haberdar edilmelidir?

I. Mesleki gelişim denildiğinde aklınıza neler geliyor?

(Mesleki gelişimi nasıl tanımlarsınız?)

Sonda: Mesleki gelişim; mesleğinizle ilgili gelişmeleri takip etmek, sorunları keşfetmek, etkili öğrenme ortamları oluşturmak, gerçekleştirdiğiniz öğretimi değerlendirmek ve iyileştirmek, alanınızda daha fazla bilgi sahibi olmak vb.

J1. Mesleki gelişiminize katkı sağlamak amacıyla hangi eğitimlerin gerçekleştirileceğine kurum nasıl karar veriyor?

J2. Bu karar verme sürecinin olumlu yanları nelerdir? Olumsuz yanları nelerdir?

J3. Öğretmenlerin ihtiyaçları nasıl belirlenmelidir?

K1. Mesleki gelişiminizi sağlamak için kurum tarafından nasıl destekleniyorsunuz? Siz kendi mesleki gelişim için neler yapıyorsunuz?

Sonda: Hizmet içi eğitimler, özel kurslar, seminerler, belli bir plan program kapsamında örneğin, öğretimi iyileştirme planı, kalkınma planı

K2. Bu eğitimleri nasıl değerlendiriyorsunuz? Bu eğitimlerin olumlu yanları nelerdir? Olumsuz yanları nelerdir?

(Bu eğitimler, sizin eksik ve zayıf yönlerinizi nasıl geliştiriyor? Öğrendiklerinizi mesleki yaşantınızda nasıl uyguluyorsunuz?)

K3. Öğretmenlerin mesleki gelişimlerini desteklemek amacıyla neler yapılmalıdır? Öğretmenler neler yapabilir?

L1. Kurumunuz bu eğitimlerin sonrasında sizlerin ihtiyaçlarının giderilip giderilmediğini nasıl belirliyor?

L2. Bu ihtiyaçların giderilip giderilmediğinden emin olmak için neler yapılmalıdır?

Öğretmen değerlendirme hakkında söylemek istediğiniz başka görüş ve öneriniz var mı?

B. PRINCIPALS INTERVIEW SCHEDULE

YÖNETİCİ GÖRÜŞME FORMU

Tarih:

Katılımcı:

Süre:

Sayın ...

Orta Doğu Teknik Üniversitesi Eğitim Bilimleri Bölümü, Eğitim Programları ve Öğretimi programında doktora öğrencisiyim. Doktora tezimi bu okulda görev alan öğretmenlerin, yöneticilerin ve uzmanların, öğretmen değerlendirme hakkındaki görüşlerini alarak ihtiyaçlarını anlamak, önerileri ve katılımları doğrultusunda öğretmenlerin mesleki gelişimlerini destekleyecek bir değerlendirme modeli hazırlamak için yürütmekteyim. Bu bağlamda, siz değerli yöneticilerin görüşlerini almak benim için çok önemli.

Bu çalışmaya katıldığınız ve zaman ayırdığınız için çok teşekkür ederim. Bu röportaj, kurumunuzda etkili bir öğretmen değerlendirme modeli oluşturma süreci hakkındaki görüşlerinizi paylaşmak için iyi bir fırsat olabilir.

Görüşmenin genel özellikleri ve gizliliği hakkında bazı hatırlatmalar yapmak istiyorum:

- Bu konuşma sadece araştırma amaçlı kullanılacaktır. Adınız hiçbir rapor vb. belgede kullanılmayacak ve araştırmacı tarafından gizli tutulacaktır.*
- Görüşmedeki ayrıntıları ve önemli bir şeyi kaçırmamam için kaydetmek istiyorum. Kayıtlar tamamen gizli tutulacaktır.*
- Konuşmamızın hiçbir bölümünde, sizi yanıltacak veya size zarar verebilecek herhangi bir şey yoktur. Bununla birlikte, görüşmenin herhangi bir aşamasında, istemiyorsanız, mülakat derhal feshedilecek ve tüm kayıt iptal edilecek.*

Görüşme yaklaşık 40 dakika sürecektir. Başlamadan önce sormak istediğiniz bir şey var mı?

Tanımlayıcı Bilgiler

- a. Kaç yıldır yöneticilik yapıyorsunuz?
- b. Bu okulda kaç yıldır çalışıyorsunuz? Daha önce hangi kurumda çalıştınız?
- c. Daha önce kaç yıl öğretmenlik yaptınız?
- d. Öğrenim düzeyinizden bahseder misiniz?
- e. Öğretmen olarak çalıştığınız dönemde bu okulda ya da daha önce çalıştığınız kurumlarda bir değerlendirme sürecinden geçtiniz mi?
- f. Bir öğretmen değerlendirme sürecinde değerlendirici olarak rol aldınız mı?

Görüşme Soruları

1. Öğretmen değerlendirme modeli denildiğinde aklınıza neler geliyor?
(Bir öğretmen değerlendirme modelini nasıl tanımlarsınız?)

2. Okulunuzda yapılan değerlendirme süreçleri ile ilgili örnekler verebilir misiniz?
Yapılan değerlendirmeler ile ilgili bir deneyiminizi paylaşır mısınız?

3. Okulunuzda yürütülen öğretmen değerlendirme süreci ve bu sürecin nasıl olması gerektiği hakkında biraz konuşalım.

A1. Öğretmenleri kimler değerlendiriyor? Değerlendirmeyi yapan kişilerin bu alandaki yeterlikleri hakkında neler düşünüyorsunuz?

Sonda: Yöneticiler, veliler, öğrenciler, aynı branştaki öğretmenler, farklı branştaki öğretmenler, kurum dışından kişiler vb.

A2. Bu kişiler tarafından değerlendirilmenin, değerlendirme sürecine olumlu etkileri nelerdir? Olumsuz etkileri nelerdir?

A3. Öğretmenleri kimler değerlendirmelidir? Değerlendirmeyi yapan kişiler hangi yeterliklere ve özelliklere sahip olmalıdır?

B1. Değerlendirmeler ne sıklıkta yapılıyor? Ne zaman yapılıyor?

B2. Değerlendirme sıklığının olumlu yönleri nelerdir? Olumsuz yönleri nelerdir?

B3. Değerlendirmeler ne sıklıkta yapılmalıdır? Ne zaman yapılmalı?

C1. Okulunuzdaki deęerlendirmede ne tr deęerlendirme araları kullanılıyor?
Sonda: Gzlem, grřme, z-deęerlendirme, uzmanlar tarafından hazırlanmıř standart formlar, anketler vb.

C2. Kullanılan bu araların, deęerlendirme srecine olumlu katkıları nelerdir?
Bu araları kullanmanın srete sebep olduęu olumsuzluklar nelerdir?

C3. Deęerlendirme modelinde ne tr deęerlendirme araları kullanılmalıdır?

D1. Okulunuzdaki deęerlendirmelerde ne tr veri kaynakları kullanılıyor?

Sonda: Ders planları, derslerde kullandığınız ara ve gereler, ęrencilerin başarı gstergeleri (karne notu, sınav sonuları vb.), gzlem sonuları, grřme sonuları, z-deęerlendirme sonuları vb.

D2. Bu veri kaynaklarının kullanılmasının deęerlendirme srecine olumlu etkileri nelerdir? Olumsuz etkileri nelerdir?

D3. Deęerlendirme modelinde ne tr veri kaynakları kullanılmalıdır?

E1. ęretmenlerin hangi yeterlikleri deęerlendiriliyor?

Sonda:

Derse hazırlık yapma (dersi planlama, sınıf ynetimi, kullanılan ara-gere ve malzemeler, farklılıkları dikkate alma vb.)

ęretimi gerekleřtirme (etkili ęrenme ortamı oluřturma, farklılıkları dikkate alma, st dzey dřnme becerilerini destekleme vb.)

ęretimi izleme (ęrencilerin geliřimlerini takip etme, eksikliklerini belirleme ve telafi etme)

Okul iindeki dięer sorumlulukları yerine getirme (okuldaki dięer grevleriniz, zmre iindeki grevlerini vb.)

İliřkiler (ęretmenlerle, velilerle, yneticilerle etkili iletiřim kurmak, iř birlięi iinde alıřmak, vb.)

Mesleki geliřim (eęitimlere katılma, yayınları takip etme, geliřmelerden haberdar olma vb.)

E2. Deęerlendirmenin bu yeterliklere gre yapılmasının olumlu ve olumsuz etkileri nelerdir?

E3. ęretmenlerin hangi yeterlikleri deęerlendirilmelidir?

F1. Okulunuzda yrtlen deęerlendirmelerde objektiflik (tarafsızlık) nasıl saęlanıyor?

(Deęerlendirme sonularına gveniyor musunuz?)

F2. Değerlendirmelerin objektifliği (tarafsızlığı) nasıl sağlanmalıdır?

G1. Değerlendirme sonuçları ne amaçla kullanılıyor?

(Öğretmenler ne amaçla değerlendiriliyor?)

Sonda: Kuruma ilişkin raporlamalar (faaliyet, memnuniyet, gelişim vb.) yapma, güçlü ya da geliştirilmesi gereken yönleriniz belirlemek, uyarı ya da ceza vermek, öğrenci başarısını arttırmak, profesyonel gelişimizi desteklemek)

G2. Değerlendirme sonuçlarının bu amaçlarla kullanılmasının olumlu yönleri nelerdir? Olumsuzlukları nelerdir?

G3. Değerlendirmeden elde edilen veriler ne amaçla kullanılmalıdır?

(Bu verilere dayalı olarak neler yapılmalıdır ve ne tür kararlar alınmalıdır?)

H1. Öğretmenler değerlendirme sonuçlarından nasıl haberdar ediliyor? Ne sıklıkla haberdar ediliyor?

Sonda: Yüz yüze görüşmelerle, yazılı iletişim yoluyla, telefonla, düzenli olarak, rastgele

H2. Bu şekilde haberdar edilmenin olumlu yönleri nelerdir? Olumsuz yönleri nelerdir?

H3. Değerlendirme sonuçları nasıl haberdar edilmelidir?

I. Öğretmenlerin Mesleki gelişim denildiğinde aklınıza neler geliyor?

(Öğretmen mesleki gelişimini nasıl tanımlarsınız?)

Sonda: Mesleki gelişim; mesleğinizle ilgili gelişmeleri takip etmek, sorunları keşfetmek, etkili öğrenme ortamları oluşturmak, gerçekleştirdiğiniz öğretimi değerlendirmek ve iyileştirmek, alanınızda daha fazla bilgi sahibi olmak vb.

J1. Öğretmenlerin mesleki gelişimine katkı sağlamak amacıyla hangi eğitimlerin gerçekleştirileceğine nasıl karar veriliyor?

J2. Bu karar verme sürecinin olumlu yanları nelerdir? Olumsuz yanları nelerdir?

J3. Öğretmenlerin ihtiyaçları nasıl belirlenmelidir?

K1. Öğretmenlerin mesleki gelişimi kurum tarafından nasıl destekleniyor?

Sonda: Hizmet içi eğitimler, özel kurslar, seminerler, belli bir plan program kapsamında örneğin, öğretimi iyileştirme planı, kalkınma planı

K2. Bu eğitimleri nasıl değerlendiriyorsunuz? Bu eğitimlerin olumlu yanları nelerdir? Olumsuz yanları nelerdir?

(Bu eğitimler, öğretmenlerin eksik ve zayıf yönlerinizi nasıl geliştiriyor? Öğretmenler öğrendiklerinizi mesleki yaşantınızda nasıl uyguluyorsunuz?)

K3. Öğretmenlerin mesleki gelişimlerini desteklemek amacıyla neler yapılmalıdır? Öğretmenler neler yapabilir?

L1. Kurumunuz bu eğitimlerin sonrasında öğretmenlerin ihtiyaçlarının giderilip giderilmediğini nasıl belirliyor?

L2. Bu ihtiyaçların giderilip giderilmediğinden emin olmak için neler yapılmalıdır?

Öğretmen değerlendirme hakkında söylemek istediğiniz başka görüş ve öneriniz var mı?

C. EXPERT INTERVIEW SCHEDULE

UZMAN GÖRÜŞME FORMU

Tarih:

Katılımcı:

Süre:

Sayın ...

Orta Doğu Teknik Üniversitesi Eğitim Bilimleri Bölümü, Eğitim Programları ve Öğretimi programında doktora öğrencisiyim. Doktora tezimi bu okulda görev alan öğretmenlerin, yöneticilerin ve uzmanların, öğretmen değerlendirme hakkındaki görüşlerini alarak ihtiyaçlarını anlamak, önerileri ve katılımları doğrultusunda öğretmenlerin mesleki gelişimlerini destekleyecek bir değerlendirme modeli hazırlamak için yürütmekteyim. Bu bağlamda, siz değerli uzmanların görüşlerini almak benim için çok önemli.

Bu çalışmaya katıldığınız ve zaman ayırdığınız için çok teşekkür ederim. Bu röportaj, kurumunuzda etkili bir öğretmen değerlendirme modeli oluşturma süreci hakkındaki görüşlerinizi paylaşmak için iyi bir fırsat olabilir.

Görüşmenin genel özellikleri ve gizliliği hakkında bazı hatırlatmalar yapmak istiyorum:

- Bu konuşma sadece araştırma amaçlı kullanılacaktır. Adınız hiçbir rapor vb. belgede kullanılmayacak ve araştırmacı tarafından gizli tutulacaktır.*
- Görüşmedeki ayrıntıları ve önemli bir şeyi kaçırmamam için kaydetmek istiyorum. Kayıtlar tamamen gizli tutulacaktır.*
- Konuşmamızın hiçbir bölümünde, sizi yanıltacak veya size zarar verebilecek herhangi bir şey yoktur. Bununla birlikte, görüşmenin herhangi bir aşamasında, istemiyorsanız, mülakat derhal feshedilecek ve tüm kayıt iptal edilecek.*

Görüşme yaklaşık 40 dakika sürecektir. Başlamadan önce sormak istediğiniz bir şey var mı?

Tanımlayıcı Bilgiler

- a. Kaç yıldır uzmanlık yapıyorsunuz?
- b. Bu okulda kaç yıldır çalışıyorsunuz? Daha önce hangi kurumda çalıştınız?
- c. Daha önce öğretmenlik yaptınız mı? Kaç yıl öğretmenlik yaptınız?
- d. Öğrenim düzeyinizden bahseder misiniz?
- e. Bu okulda ya da daha önce çalıştığınız kurumlarda öğretmenler değerlendirme yapıyor muydu?
- f. Siz bir öğretmen değerlendirme sürecinde değerlendirici olarak rol aldınız mı?

Görüşme Soruları

1. Öğretmen değerlendirme modeli denildiğinde aklınıza neler geliyor?
(Bir öğretmen değerlendirme modelini nasıl tanımlarsınız?)

2. Okulunuzda yapılan değerlendirme süreçleri ile ilgili örnekler verebilir misiniz?
Yapılan değerlendirmeler ile ilgili bir deneyiminizi paylaşır mısınız?

3. Okulunuzda yürütülen öğretmen değerlendirme süreci ve bu sürecin nasıl olması gerektiği hakkında biraz konuşalım.

A1. Öğretmenleri kimler değerlendiriyor? Değerlendirmeyi yapan kişilerin bu alandaki yeterlikleri hakkında neler düşünüyorsunuz?

Sonda: Yöneticiler, veliler, öğrenciler, aynı branştaki öğretmenler, farklı branştaki öğretmenler, kurum dışından kişiler vb.

A2. Bu kişiler tarafından değerlendirilmenin, değerlendirme sürecine olumlu etkileri nelerdir? Olumsuz etkileri nelerdir?

A3. Öğretmenleri kimler değerlendirmelidir? Değerlendirmeyi yapan kişiler hangi yeterliklere ve özelliklere sahip olmalıdır?

B1. Değerlendirmeler ne sıklıkta yapılıyor? Ne zaman yapılıyor?

B2. Değerlendirme sıklığının olumlu yönleri nelerdir? Olumsuz yönleri nelerdir?

B3. Değerlendirmeler ne sıklıkta yapılmalıdır? Ne zaman yapılmalı?

C1. Okulunuzdaki değerlendirmede ne tür değerlendirme araçları kullanılıyor?

Sonda: Gözlem, görüşme, öz-değerlendirme, uzmanlar tarafından hazırlanmış standart formlar, anketler vb.

C2. Kullanılan bu araçların, değerlendirme sürecine olumlu katkıları nelerdir? Bu araçları kullanmanın süreçte sebep olduğu olumsuzluklar nelerdir?

C3. Değerlendirme modelinde ne tür değerlendirme araçları kullanılmalıdır?

D1. Okulunuzdaki değerlendirmelerde ne tür veri kaynakları kullanılıyor?

Sonda: Ders planları, derslerde kullandığınız araç ve gereçler, öğrencilerin başarı göstergeleri (karne notu, sınav sonuçları vb.), gözlem sonuçları, görüşme sonuçları, öz-değerlendirme sonuçları vb.

D2. Bu veri kaynaklarının kullanılmasının değerlendirme sürecine olumlu etkileri nelerdir? Olumsuz etkileri nelerdir?

D3. Değerlendirme modelinde ne tür veri kaynakları kullanılmalıdır?

E1. Öğretmenlerin hangi yeterlikleri değerlendiriliyor?

Sonda:

Derse hazırlık yapma (dersi planlama, sınıf yönetimi, kullanılan araç-gereç ve malzemeler, farklılıkları dikkate alma vb.)

Öğretimi gerçekleştirme (etkili öğrenme ortamı oluşturma, farklılıkları dikkate alma, üst düzey düşünme becerilerini destekleme vb.)

Öğretimi izleme (öğrencilerin gelişimlerini takip etme, eksikliklerini belirleme ve telafi etme)

Okul içindeki diğer sorumlulukları yerine getirme (okuldaki diğer görevleriniz, zümre içindeki görevlerini vb.)

İlişkiler (öğretmenlerle, velilerle, yöneticilerle etkili iletişim kurmak, iş birliği içinde çalışmak, vb.)

Mesleki gelişim (eğitime katılma, yayınları takip etme, gelişmelerden haberdar olma vb.)

E2. Değerlendirmenin bu yeterliklere göre yapılmasının olumlu ve olumsuz etkileri nelerdir?

E3. Öğretmenlerin hangi yeterlikleri değerlendirilmelidir?

F1. Okulunuzda yürütülen değerlendirmelerde objektiflik (tarafsızlık) nasıl sağlanıyor?

(Değerlendirme sonuçlarına güveniyor musunuz?)

F2. Değerlendirmelerin objektifliği (tarafsızlığı) nasıl sağlanmalıdır?

G1. Değerlendirme sonuçları ne amaçla kullanılıyor?

(Öğretmenler ne amaçla değerlendiriliyor?)

Sonda: Kuruma ilişkin raporlamalar (faaliyet, memnuniyet, gelişim vb.) yapma, güçlü ya da geliştirilmesi gereken yönleriniz belirlemek, uyarı ya da ceza vermek, öğrenci başarısını arttırmak, profesyonel gelişimizi desteklemek)

G2. Değerlendirme sonuçlarının bu amaçlarla kullanılmasının olumlu yönleri nelerdir? Olumsuzlukları nelerdir?

G3. Değerlendirmeden elde edilen veriler ne amaçla kullanılmalıdır?
(Bu verilere dayalı olarak neler yapılmalıdır ve ne tür kararlar alınmalıdır?)

H1. Öğretmenler değerlendirme sonuçlarından nasıl haberdar ediliyor? Ne sıklıkla haberdar ediliyor?

Sonda: Yüz yüze görüşmelerle, yazılı iletişim yoluyla, telefonla, düzenli olarak, rastgele

H2. Bu şekilde haberdar edilmenin olumlu yönleri nelerdir? Olumsuz yönleri nelerdir?

H3. Değerlendirme sonuçları nasıl haberdar edilmelidir?

I. Öğretmenlerin mesleki gelişim denildiğinde aklınıza neler geliyor?
(Öğretmen mesleki gelişimini nasıl tanımlarsınız?)

Sonda: Mesleki gelişim; mesleğinizle ilgili gelişmeleri takip etmek, sorunları keşfetmek, etkili öğrenme ortamları oluşturmak, gerçekleştirdiğiniz öğretimi değerlendirmek ve iyileştirmek, alanınızda daha fazla bilgi sahibi olmak vb.

J1. Öğretmenlerin mesleki gelişimine katkı sağlamak amacıyla hangi eğitimlerin gerçekleştirileceğine nasıl karar veriliyor?

J2. Bu karar verme sürecinin olumlu yanları nelerdir? Olumsuz yanları nelerdir?

J3. Öğretmenlerin ihtiyaçları nasıl belirlenmelidir?

K1. Öğretmenlerin mesleki gelişimi kurum tarafından nasıl destekleniyor?

Sonda: Hizmet içi eğitimler, özel kurslar, seminerler, belli bir plan program kapsamında örneğin, öğretimi iyileştirme planı, kalkınma planı

K2. Bu eğitimleri nasıl değerlendiriyorsunuz? Bu eğitimlerin olumlu yanları nelerdir? Olumsuz yanları nelerdir?

(Bu eğitimler, öğretmenlerin eksik ve zayıf yönlerinizi nasıl geliştiriyor?

Öğretmenler öğrendiklerinizi mesleki yaşantınızda nasıl uyguluyorsunuz?)

K3. Öğretmenlerin mesleki gelişimlerini desteklemek amacıyla neler yapılmalıdır? Öğretmenler neler yapabilir?

L1. Kurumunuz bu eğitimlerin sonrasında öğretmenlerin ihtiyaçlarının giderilip giderilmediğini nasıl belirliyor?

L2. Bu ihtiyaçların giderilip giderilmediğinden emin olmak için neler yapılmalıdır?

Öğretmen değerlendirme hakkında söylemek istediğiniz başka görüş ve öneriniz var mı?

D. FOCUS GROUP INTERVIEW SCHEDULE

A. AÇILIŞ

Merhaba,

Bu odak grup görüşmesinin amacı daha önce sizlerle yapılan görüşme sonuçları ile oluşturulan öğretmen değerlendirme modelinin her bir ögesi ile ilgili sizlerin görüşlerini almaktır.

Görüşmeye başlamadan önce bu çalışmaya katıldığınız ve zaman ayırdığınız için çok teşekkür ederim. Görüşmenin genel özellikleri ve gizliliği hakkında bazı hatırlatmalar yapmak istiyorum:

- *Bu konuşma sadece araştırma amaçlı kullanılacaktır. Adınız hiçbir rapor vb. belgede kullanılmayacak ve araştırmacı tarafından gizli tutulacaktır.*
- *Görüşmedeki ayrıntıları ve önemli bir şeyi kaçırmamam için kaydetmek istiyorum. Kayıtlar tamamen gizli tutulacaktır.*
- *Konuşmamızın hiçbir bölümünde, sizi yanıltacak veya size zarar verebilecek herhangi bir şey yoktur. Bununla birlikte, görüşmenin herhangi bir aşamasında, istemiyorsanız, mülakat derhal feshedilecek ve tüm kayıt iptal edilecek.*

Görüşme yaklaşık 1 saat sürecektir. Başlamadan önce sormak istediğiniz bir şey var mı?

B. TANITMA

Sizlerden kısaca kendinizi tanıtmınızı rica edeceğim. Bu amaçla sizlere dağıtmış olduğum forma adınızı, soyadınızı, öğrenim düzeyinizi, kaç yıldır öğretmenlik yaptığınızı ve bu kurumda kaç yıldır çalıştığınızı yazar mısınız?

Bu model oluşturulurken 20 öğretmen, 9 yönetici ve 7 uzmandan görüş toplandı. Görüşler incelendiğinde çoğunluğun ortak belirttiği hususlar doğrultusunda bu model planlandı. Sizlere planlan modeli kısaca tanıtmak isterim. Bu modelin amacı öğretmen yeterliklerini değerlendirmek ve değerlendirme sürecine bağlı olan öğretmen mesleki gelişimi için sistematik destek sağlamaktır... (Bu bölümde önerilen model sunum yardımıyla tanıtılır.)

C. SORULAR

1. Sizlerden gelen görüşler doğrultusunda modelin işlenişinde ilk sayfada bahsedilen hususlara yer verilecektir. Bu hususların her biri ile ilgili ayrı ayrı görüşlerinizi alabilir miyim?

2. Sizlerle yapılan görüşmelerde bu model kapsamında hangi öğretmen yeterliklerinin değerlendirilmesi gerektiğini konuşmuştuk. Görüşmelerden elde edilen sonuçlar doğrultusunda belirtilen yeterlik alanlarının yer alması uygun görülmüştü. Bu yeterlik alanları hakkında görüşlerinizi alabilir miyim?

- a) Bu yeterlik alanlarına ekleme yapmak ister misiniz? Açıklayın.
- b) Bahsettiğiniz alanlar bahsedilen beş yeterlik alanına dâhil edilebilir mi? Nasıl?
- c) Bu alanlardan hangisi çıkartılmalıdır? Neden?
- d) Bu yeterlik alanları arasında önem sırası olmalı mıdır? Bu kurum için düşündüğünüzde hangi yeterlik alanının değerlendirilmesi diğerinden daha önemlidir?

3. Bu modeldeki yeterlik ölçütleri hem okulunuzda yeni başlayan hem de tecrübeli öğretmenler için aynı mı olmalıdır?

- a) Değilse hangi ölçütler farklılık göstermelidir? Neden?
- b) Yeterlik alanlarının önem sırası yeni başlayan ve tecrübeli öğretmen için farklılık göstermeli midir? Nasıl bir ağırlık andırma önerirsiniz?

4. Belirtilen yeterlik ölçütleri kim tarafından, nasıl ve öğretim yılının hangi zaman diliminde değerlendirileceği belirtilmiştir. Tablodaki her bir sütunu tek tek inceleyelim (Katılımcıların inceleyebilmesi için 5 dk süre verilir) Her bir sütun için.

- a) Belirtilen ölçütler bu kişi ya da kişiler tarafından değerlendirilebilir mi? Değerlendirilemezse nedenlerini açıklayınız? Bu ölçüt başka kimler tarafından değerlendirilebilir? Gözlem yapan kişilerin objektif olamayacağından bahsetmişsiniz bu durumda gözleme aynı anda birden fazla kişinin girmesi çözüm olabilir mi? Nasıl?
- b) Belirtilen ölçütler için kullanılması planlanan yöntem ve veri toplama aracı uygun mudur? Değilse nedenlerini açıklayınız? Bu ölçütler başka nasıl değerlendirilebilir?
- c) Belirtilen ölçütler için planlanan zaman dilimi uygun mudur? Değilse nedenlerini açıklayınız? Bu ölçütler başka ne zaman değerlendirilebilir?
- d) Yapılan görüşmeler incelendiğinde öğretmenlerin çoğunluğunun ders planlarının değerlendirilmesinin süreci yansıtmayacağı için yer almaması gerektiğini ifade etmişti. Böyle bir model dâhilinde ders planının kullanılması konusunda ne düşünüyorsunuz?

5. Bu modelde her dönemin sonunda müdür ve öğretmenin bir araya gelerek tüm değerlendiricilerden elde edilen raporlar yardımıyla bir gelişim planı hazırlanması ve

öğretmenin bu gelişim planı doğrultusunda kendi mesleki gelişiminin desteklenmesi önerilmektedir. Bu performans değerlendirme görüşmesi ile ilgili görüşleriniz nelerdir?

a) Öğretmenlerin mesleki gelişimini destekleyen bir model bu modelden farklı olarak nasıl tasarlanabilir?

6. Bu modelin uygulanabilirliği hakkında ne düşünüyorsunuz? Siz böyle bir model ile değerlendirilmek ve/veya değerlendirme yapmak ister miydiniz? Neden/ neden değil?

D. KAPANIŞ

Katılımlarınız için çok teşekkür ederim. Bundan sonraki aşamada sizlerin karar verdiği yeterlik alanlarına ve değerlendirme yöntemlerine bağlı olarak ölçeklerde yer alamsı düşünülen ölçütler liste halinde sizlerle paylaşılacak ve bu ölçütler ile ilgili sizlerin görüşleri alınacaktır.

**E. THE CODEBOOK OF 1ST DATA COLLECTION PROCESS
(INDIVIDUAL INTERVIEWS)**

RQ1: What kind of teacher evaluation practices are carried out for classroom teachers at this private school?		
Theme 1. Evaluation Process		
Code	Description	Example
No evaluation schedule or pattern	Statements showing that participants were not evaluated at a certain time or at certain intervals.	"Evaluations are not made in a certain time period, so it can happen at any time."
No standardized forms	Explanations that the evaluations, especially the observation done by the principal, were made only by taking notes without a form.	"I do not think that the evaluations are made with a certain form. Our principal just sat in the classroom and took notes."
No written purpose for evaluation	The state of being unaware of the existence of an evaluation purpose or stating that the purpose of the evaluation was not informed in writing.	"I do not know clearly, what the evaluation was made for, and suppose the evaluation was conducted based on a purpose. Unfortunately, no one informed us about this purpose written or orally."
No defined standards or criteria	Statements showing that the evaluation was not made in accordance with certain criteria or standards.	"I do not know what criteria are used. I don't even think the criteria exist."
Data source: Average achievement scores	Reporting the results of school-wide achievement tests a using them for teacher evaluation or regarding the test scores as equivalent to the teacher's success.	"After each exam, the average score of the class is used to evaluate the teacher."
Data source: Lesson plans	Use of lesson plans which were checked by assistant principals for evaluation purposes.	"I send lesson plans to the assistant principal every week. They also evaluate me through these plans."
Data source: Informal principal observation	Informal classroom observation process carried by the school principal.	"This year, our school principal came to the classroom and observed during a lesson by taking notes."
Other data source: Opinions of colleagues and parents	Seeking the opinions of the head of the department or the parents for the purpose of evaluation.	"I think that the opinions of our head of department are used in the evaluation." "Since we are a private school, the opinions of the parents are also included in the evaluation."

Theme 2. Use of Evaluation Data		
Code	Description	Example
Improving student achievement	The use of assessments to increase student achievement, even if not explicitly stated.	"It is not clearly stated, but I think the only purpose is to increase student success."
Predicting teachers' success	The use of assessments, especially exam results, to predict the teacher level of success.	"According to the exam results, the teacher is considered as successful or unsuccessful."
Improving the quality of school	Implying that evaluations may indirectly affect the quality of the school.	"As student success increases, so does the quality of the school."
Providing summative feedback	Giving result-oriented feedback on general issues at a meeting attended by teachers teaching at the same grade level.	"At the end of the semester, general information was given to all teachers."
Providing individual formative feedbacks	One-on-one meetings with teachers only in problematic or undesirable situations	"If there is a complaint from the parent, we receive one-to-one feedback."

RQ2: How are teacher evaluation practices perceived in terms of strengths and weaknesses by classroom teachers, principals, and experts at this private school?		
Theme 1. Propriety		
Code	Description	Example
Not providing policies and procedures	Problems arising from the absence of a policy or procedure that states the purpose, process, schedule, or timeline of the evaluation. Problems that arise due to the fact that the evaluations are far from being systematic.	"The purpose, and the process is unclear. This brings the question of why I'm being evaluated." "The lack of a certain frequency of observation means that everyone should be as comfortable as possible."
No access to evaluation results	Problems caused by not sharing evaluation or observation results. Not giving feedback to the teachers.	"I haven't received any notification or feedback about what happened in the evaluation yet, for what purpose this evaluation result was used."
No balanced evaluation	Receiving notifications or feedbacks only for the weaknesses. Receiving one-to-one feedback, especially when they were not successful in the classroom. Not sharing the aspects in which teachers are effective or good at.	"Principals only talk about my weaknesses when they want to speak individually. I never heard a good thing about my teaching"
Theme 2. Utility		
Code	Description	Example
Lack of explicit criteria	The absence of criteria, standards or indicators negatively affects the usefulness of the results or defensibility of the judgment.	"Principals has opinions about me, but I do not know how they have achieved this view. I don't know which performance indicator was used. [...] I can neither defend myself," "I think it is meaningless

		and not useful when the evaluation is not done with the criteria.”
Lack of functional reporting	Failure to record and report evaluation results and feedback to guide teacher development.	“If one had checked me and kept the report, it would keep me vigorous. In this way, I know my weaknesses and follow developments.”
Theme 3. Accuracy/ Misinterpretation of teacher effectiveness		
Code	Description	Example
Competitive environment	Inaccurate measurement of effectiveness by comparing teachers with each other, creating a competitive environment among them	“Everyone is compared to each other. Being successful is somewhat like the ego war between teachers. If you have achieved significant success in your class, this is the success of your class. Should not be compared to others.”
Negative aspects of the observation process	The fact that the teacher behaves differently from the normal due to the anxiety caused by the observation process causes a mistake in the evaluation.	“We constantly think that these observation results will be used against us. For that reason, perhaps we are trying to explain better or being more positive when we are reacting to the student”
Theme 3. Accuracy/ Biased identification		
Code	Description	Example
Biased judgements based on experience	Being evaluated differently according to teaching experiences and evaluations are made in favor of experienced teachers.	“[...] experienced teachers continue to practice what they knew twenty years ago, and I have never seen that they have received any warning”
Biased judgements based on personal relations	The situation where the friendship relations between individuals prevent negative criticism of each other or the negative relations between individuals cause biased evaluations.	“If I have to give an example from myself, I do not want to evaluate a friend I love negatively.” I'm not good with some teachers. She can make observations just to evaluate badly.”
Theme 3. Accuracy/ Reliable Information		
Code	Description	Example
Ignoring other important domains of learning	Equating student achievement with teacher success tends to ignore teacher efforts to improve some important areas such as the affective domain.	“I don't think it's right to use student achievement as a single source. No one knows what I do for affective domain. Then it was up to my conscience to develop this domain.”
Not reflecting the complexity of teaching/learning process	The fact that student success is equated with teacher success and using lesson plans as the only data source causes the complexity of the process carried out by the teacher in the classroom to be ignored.	“Only the average test score cannot be my success. I think the effort I put into the classroom is more meaningful.” “Lesson plans do not reflect the classroom environment or my teaching, which we often fail to follow. Teaching is more difficult and complex than this plan.”

RQ3: What are the recommendations of classroom teachers, principals, and experts of this school to develop an effective teacher evaluation model?

Theme 1. Evaluation Process/ Purpose of Evaluation

Code	Description	Example
Evaluating and developing teacher qualifications	Statements that the purpose of evaluation should be to evaluate the qualifications of the teacher.	“Evaluation should be for the development of teachers and their qualifications”
Improving student’s achievement	Statements that the purpose of the evaluation may be to increase student achievement if the average scores obtained from school-wide exams are not considered as the only data source and/or if the teacher's contribution to the individual success of the students is correctly evaluated.	“Success is important, but I don't find it right to use student success as a single source” “We are teachers; of course, we do our best to increase the success of the student. But rather than the success of the whole class, students' individual achievement should be looked at to understand what we have contributed to this student individually.”

Theme 1. Evaluation Process/ Teacher Qualification to be assessed

Code	Description	Example
Communication and Collaboration	Statements explaining the importance of assessing the ability to communicate effectively and positively with other teachers, principals, and students and to work collaboratively with colleagues	“Everyone has to work together and collaboratively, whether it is related to classes or extracurricular activities. An indispensable part of cooperation is to use communication effectively. That is, teachers need improvement and change in the field of communication and collaboration”
Instruction	Situations regarding the necessity of including effective teaching among the competence areas to be evaluated	“Perhaps the most important evaluation of what happens in the classroom, namely instruction.”
Service to the school	Statements explaining the importance of assessing service to school including professional and in-school responsibilities of teachers such as making contributions to the development of the school, participating in the activities carried out in the school, keeping duties in the school, and taking responsibility in ceremonies.	“In this school, it is very important to evaluate the responsibilities within the school, keeping duties or, for example, the behavior in the ceremonies.”
Planning and Preparation	Situations that reveal that the process of lesson preparation or prepared lesson plan should be evaluated together with the observation process of the instruction.	“One of the most critical factors for the success of the course is the well-structured preparation phase. Before observing, teachers may be asked what kind of preparations or planning we are doing.”
Monitoring and managing learning	Opinions supporting that following the cognitive or affective development of students and taking precautions when necessary are among the qualifications that should be evaluated.	“A good teacher should already follow the academic and affective development of his student. The classroom teacher knows students ‘shortcomings best. I think this competence is important.”

Professional development	Opinions that it is important for teachers to develop themselves by receiving necessary training and that professional development competence can be evaluated for this purpose.	“Professional development should be evaluated, I think, as an essential qualification. Participation in training is critical, and I think this is an important criterion to be evaluated.”
Theme 1. Evaluation Process/ Evaluation Schedule		
Code	Description	Example
Except beginning of each academic period	Statements that it would not be appropriate to make assessments and classroom observations as soon as the academic term starts.	“At the beginning of the semester, we are dealing with the adaptation of students a lot, I think there should be no observations at these times.”
One or two observations for each month	Opinions that it would be appropriate to make at most one or two observations per month.	“It is sufficient to be observed at most two times in a month. Too much will cause the lesson to be divided.”
Three classroom observations for each semester	Opinions that it would be appropriate to make at most three observations in each semester	“There may be a total of 2 or 3 observations during the period.”
Theme 1. Evaluation Process/Evaluator Features		
Code	Description	Example
Objective/Fair	Emphasizing the importance of the evaluator being objective and/or fair for the effectiveness of the evaluation.	“Perhaps the essential feature of an observer is objectivity the most important and possibly the most difficult.”
Positive/effective communication skills	The importance of the evaluator having effective and positive communication skills for the efficiency of the evaluation.	“The evaluator must establish positive communication to create an environment where I can internalize his criticism of me.”
Has teaching experience	The importance of the evaluator having teaching experience and knowing the classroom environment for the assessment to be accurate.	“[...] it is crucial to have a person who knows primary school students and even works as a primary school teacher”
Empathy skills	The importance of making evaluations by people who can empathize and consider the situation in the classroom.	“If she cannot put herself in my place, she cannot correctly assess what I am doing in the classroom, and it incorrectly determines where I am missing.”
Enough knowledge about school	The fact that the evaluator works at the school and/or knows the school is important in evaluating the teacher working at that school	“A person outside the school could not make objective evaluations more than the evaluations made by the people who knew the school”
Evaluation knowledge and skills	Evaluator should have qualifications and skills required, such as using the tools effectively, fulfilling the requirements of the observer role, and making accurate evaluations	“Must be knowledgeable in making observations and conducting interviews. If possible, should have evaluated in this way before.”
Theme 1. Evaluation Process/ Enabling Evaluation Dynamics		
Code	Description	Example
Training schedule for evaluators	The importance of evaluators receiving training on knowledge and skills such as the purpose of the model, using data collection tools adequately, effective	“Evaluators may not have much information and skills about the evaluation, but they can be competent in this field with the training they received.”

	implementation of the evaluation process, and giving correct feedback.	
Information meetings for evaluated person	The importance of informing the evaluated person about the purpose of the model, the use of the tools, how the process will be carried out and how the result will be informed.	“There should be training at the very least, including the purpose of this evaluation model, data collection tools to be used, and even observation.”
Theme 1. Evaluation Process/ Data source		
Code	Description	Example
Principals	Evaluation of teachers by school principals.	“I really want the principal to observe my class.”
Classroom teachers	Evaluation of teachers by teacher another classroom teacher.	“I am sure a classroom teacher can observe me very well in the classroom.”
Self	Evaluation of teachers by teacher herself.	“Perhaps the most important thing is that the teacher can evaluate herself. Because a teacher knows herself very well.”
Math/science teachers	Evaluation of teachers by math and science teachers.	“Teachers from different branches can say something about the field. It is helpful to be aware of new information in this field, especially in science and mathematics lessons, or give me feedback if I am wrong or missing.”
School Experts	Evaluation of teachers by school experts.	“Expert working in this school could also evaluate teachers, because they are qualified to evaluate the teachers.”
Students	Evaluation of teachers by students.	“A student going to 4th grade may evaluate his teacher.”
Parents	Evaluation of teachers by parents.	“I have a parent who is an educator. Her feedback is important. For example, she can evaluate.”
Theme 1. Evaluation Process/ Evaluation Method		
Code	Description	Example
Observations	Evaluation of the teacher through multiple classroom observations.	“Making multiple observations is the only way to understand the teacher performance and classroom practices totally.”
Debriefing after observation	Conducting interviews after the observation for debriefing to reduce the misunderstandings caused by the observer in the classroom.	“If the observer talks to me about what is observed in the classroom, maybe I can clear up any misunderstandings about what is happening in the classroom at that moment.”
Self-evaluation	Evaluation of the teacher through self-evaluation and/or self-evaluation forms. .	“But I think as a teacher, we must first evaluate ourselves and get used to these forms.”
Theme 2. Use of Evaluation Data/ Access to Evaluation Information		
Code	Description	Example
Importance of confidentiality	Sharing the results, opinions or decisions obtained from the	“No one should know my result and only I should be able to reach it.”

	evaluation only with the teacher, keeping confidential and accessing the information only by the teacher.	
Individual feedback (face to face)	Individual reporting of the results, opinions or decisions obtained from the evaluation	“Collective feedback has no effect. One-on-one and individual feedback should be given.”
Theme 2. Use of Evaluation Data/ Propriety		
Code	Description	Example
Balanced Evaluation (identifies both strengths and weaknesses)	Individual-oriented feedback on both teacher strengths and weaknesses to support professional development.	“In fact, when explaining the evaluation results, first the strengths and then the shortcomings can be explained. Weaknesses should not only be justified but also how these weaknesses can turn into strengths should be described.”
Professional interactions	The necessity of reporting feedback or work that needs to be done through respectful and positive communication	“We should always be moderate, courteous and respectful when giving feedback to the teacher about her work.”
Theme 2. Use of Evaluation Data/Utility		
Code	Description	Example
Explicit criteria for usefulness of the results or defensibility	Evaluations should be on explicit criteria so that interpretation and judgment based on these criteria make sense and, thus, an open and defensible assessment environment is created.	“Evaluations should be criteria-based, and the results should inspire me. So, I should be able to develop it for myself with the help of these results. The criteria ultimately give clear information on what I should improve on performance.”
Functional reporting	The necessity of a report including feedbacks and teacher growth plan.	“A detailed report with feedback and you can follow my progress would be very useful.”
One written report per year	Reports should be given at the end of each education period to give teachers enough time for their development.	“It would be unfair to wait until the end of the year for students in this class. The teacher should get feedback as soon as possible about this. Feedback should be given at the end of the first semester at the latest.”
Theme 2. Use of Evaluation Data/Accuracy		
Code	Description	Example
Bias Identification and Management	Suggesting that people with such personal relationships should not observe each other so that the good or bad individual relations between teachers do not prevent objective evaluations.	“A close friend of mine cannot observe me objectively. It would be better for a teacher with whom I am not very familiar to observe.”
Equal evaluation regardless of teaching experience	The importance of evaluating all teachers in the same way and with the same criteria, regardless of their professional experience.	“Therefore, assessment should be the same for all, irrespective of the teaching experience, and even the criteria should be the same.”
Use of multiple data source and method	The necessity of multiple data collection process to provide reliable information and to be accurate in evaluation.	“My teaching should not be evaluated by looking at a single performance. In fact, it should not be done only an observation. [...] Observation, interview all must be used. It can be a structured observation form. Self-

		assessment should also be linked to other observations.”
Open communication	An environment of trust that gives the teacher the opportunity to express and/or defend himself without fear in the process of giving feedback or creating a professional development plan.	“Evaluation results should be explained face to face, and the teacher should defend herself when necessary. The teacher should be allowed to protect her right and express herself. She should not feel fear to express herself.”
Positive communication	The importance of establishing positive communication with the teacher in the process of giving feedback or creating a professional development plan.	“I can express myself more comfortably in a different, more positive environment.”
RQ4: What kind of professional development processes are carried out at this private school and what is needed?		
Theme 1. Training planning		
Code	Description	Example
Needs of most teachers	Preferring the topics reported by most of the teachers in planning the trainings targeting professional development	“The most-posted title is selected. Then, finally, a seminar on that topic is given.”
Based on educational issues on the agenda	Preferring topics or people that are on the agenda and/or popular in planning trainings targeting professional development	“There are things like that, for example, a person is trendy at that time, writes about education, and can come to our school and give a seminar.”
Opinions of school experts	Considering the suggestions of the experts working at the school in planning the trainings targeting professional development.	“These experts are constantly working with teachers. I know they are sometimes asked to determine the training topics.”
Theme 2. Strengths of professional development process		
Code	Description	Example
Supportive environment	The institution's support for professional development and providing opportunities for training.	“For example, I want to attend many pieces of training, and the school board says, let me know the training you want to join. In fact, the school supports us [...]”
Theme 3. Problems of professional development process		
Code	Description	Example
Trainings in large groups	The problem of the ineffectiveness of the general trainings given to the large masses since they do not meet individual needs, what is learned is not permanent, and the participants cannot actively participate.	“Why do we sit in a big room that is very crowded and listen for hours in training? I honestly remember very little of what I learned when I left training.”
Failure to evaluate the impact of trainings	Problems arising from not measuring the applicability, effectiveness or permanence of the	“We received a lot of training, how many of them do we apply properly in the classroom? There are no observations about this, so it should

	training received for professional development.	be evaluated correctly. We forget most of it.”
Repetition of trainings of the same content	The problem of repetition of training stemming from the incorrect determination of teachers' professional development needs.	“You cannot give the right training unless you determine the need correctly. So, some ineffective ones are being repeated.”
Theme 4. Needs of professional development process		
Code	Description	Example
Determining the needs based on teacher evaluation	The necessity of using teacher evaluation results in determining the needs of teachers to plan and implement the professional development trainings.	“Determining the teacher's needs is, in fact, can be done through an accurate teacher evaluation system. The teacher also accepts the weaknesses as a result and will be willing to receive training.”
Measuring the effectiveness of training	The necessity of gathering data to understand the effectiveness of the trainings by conducting observations and interviews.	“Okay, I took this training, but then did I practice what I learned? It is only evident through classroom observation.”
Access to different instructional processes	Providing teacher access to different classroom practices, examples, or activities.	“I think it is essential for teachers to see the practices in each other's classrooms in terms of professional development.”

F. THE CODEBOOK OF 2ND DATA COLLECTION PROCESS (FOCUS GROUP INTERVIEWS)

RQ1: What kind of teacher evaluation practices are carried out for classroom teachers at this private school?		
Theme 1. Evaluation Process/Important Issues		
Code	Description	Example
Repetition of training	The repetition of the training about the model developed with regular intervals.	Sometimes what is said at the beginning of the academic year can be forgotten at the end of the term. I think it is important to repeat these training, that is, to make occasional reminders.
Practical training	The trainings to be given about the developed model should be practical and/or based on the application.	These observers need to practice using the form, just like piloting. These applications can also be in training, for example.
Consistency between observation processes	The importance of ensuring consistency between the observer, observation process and the observation results when more than one observation was made.	If observers use the observation form differently, there is an inconsistency here. Everyone should understand the form in the same way.
Theme 1. Evaluation Process/ Organization of the Qualifications		
Reorganizing or Changing qualifications	Reorganization or changing of dimensions and sub-dimensions of the qualifications due to reasons such as overlapping or being intertwined.	In fact, all of the mentioned areas of "relevance to the student, being effective, using resources correctly" seemed to be related to "designing teaching".
Equal importance to each qualification	Statements about the importance of each qualification area to provide a holistic evaluation process.	If a teacher does not do in-school responsibilities, this causes injustice. Therefore, it is crucial and cannot be less important than others.
Same qualifications for	All the qualifications mentioned should be the same	I think it is not a problem to evaluate inexperienced and

experienced and novice teachers	for both novice and experienced teachers.	experienced teachers within the same qualification areas. In this school, everyone comes with a specific experience [...]
Theme 1. Evaluation Process/Written Schedule of the Model		
Code	Description	Example
Classroom teacher evaluating communication and collaboration	The necessity of evaluating the competence area of working in communication and collaboration not only by the head of the department, but also by other classroom teachers.	I want my communication to be evaluated by another class teacher.
Using lesson plans as data source	The importance of evaluating lesson plans with more than one data source as in this model	If the lesson plans in this model will not be determinant alone, evaluating the plans can be used because the classroom teaching is also examined.
Theme 2. Evaluation Results/Access to Evaluation Information		
No scores	The situation where it is unnecessary to give scores in a model that supports professional development	Let's say I got 70 points, and I was successful in this evaluation. Or I got 20, and I failed. I think such a judgment has been made to find a reason, not for professional development but to enforce the teacher.
Theme 2. Evaluation Results/ Feasibility		
Willingness to use the model	Willingness to use this evaluation model that reflects teachers' needs and suggestions	I wish this model had been applied immediately in our school. I believe this model will work as it will meet the needs of most teachers.
Reducing the duties and responsibilities of the head of department	Reducing the course load of the head of the department to make these evaluations feasible	In our school, the head of the department has many duties. can do all the evaluations if these tasks are reduced.
Theme 3. Guide and Tools of the Model/ Need for a detailed guide including		
Importance of the teacher evaluation	The necessity of including the importance of teacher evaluation in the guide.	Teacher evaluation is a difficult concept to accept. In order not to be afraid of evaluation, you need to know what it means and its importance. This importance

		should definitely be included in the guide.
The model development process	The necessity and importance of including the model development process in the guide.	“How this model is developed” should be added to the guide so that everyone knows. Even a teacher who has just started the institution should know.
The purpose of the evaluation	The necessity and importance of including the purpose of the evaluation in the guide.	The fact that the purpose is in a written form makes it clear to everyone. This evaluation is not made to punish. They can easily apply it.
Detailed explanation of the qualifications	Qualification and explanation of the qualifications needed to be included in the guide in a clear and understandable way	For the effectiveness of the evaluation, it is very important to use a common language, that is, to evaluate with the same terminology. This can only be achieved by defining behaviors. What are the teacher behaviors that explain the qualification?
Use of data collection tools	The necessity and importance of including the use of data tools in the guide.	All the tools should be in the guide. They should be introduced briefly, and it can even be mentioned in the guide [...]
Consistency, objectivity, for the process	The necessity of explaining that the process of the model will be carried out consistently and objectively in the guide	If the first problem that comes to our mind here is objectivity and confidentiality, this may come to everyone's mind. We should clearly state how we will achieve this in the guide.
Openness and confidentiality for the results	The importance of explaining that the result will be shared in accordance with the principle of openness and confidentiality in the guide	What if it's not consistent? Will I be given the right to express myself? It is so important that all of them should be included in the guide one by one.
Theme 3. Guide and Tools of the Model/ Need for developed tools		
Piloting the tools	The need for piloting prepared data collection tools.	Maybe we can't evaluate some criteria. Then that criterion should be changed, for example. A pilot may be necessary.

G. THE INITIAL TEACHER EVALUATION MODEL

Modelin Amacı: Bu öğretmen değerlendirme modelinin amacı öğretmen yeterliklerini değerlendirmek ve değerlendirme sürecine bağlı olan öğretmen mesleki gelişimi için sistematik destek sağlamaktır.

Modelde Yer Alması Planlanan Hususlar

- Güvenilir bilgi için birden fazla veri kaynağı kullanma
- Farklı zamanlarda ve farklı kişiler tarafından bütüncül değerlendirme yapma
- Veri toplama araçlarının belli ölçütler doğrultusunda hazırlanması
- Değerlendirme sonucunda hem güçlü hem de zayıf yönleri yönelik dönüt verilmesi
- Yüz yüze raporlama ile gizlilik sağlanması
- Yüz yüze görüşmelerle savunulabilir ortam sağlanması
- Hem değerlendiriciler hem de değerlendirilenler için eğitimin sağlanması.

Değerlendirilecek Yeterlik Alanları ve Alt Boyutlar

A. Plan Yapma ve Derse Hazırlık

- A1. Alan ve Pedagoji Bilgisi
- A2. Öğrenci Gelişimine, İlgi ve İhtiyaçlara Uygunluk
- A3. Program Okuryazarlığı
- A4. Bilginin etkin bir şekilde sunulması
- A5. Kaynakları etkili kullanma
- A6. Öğrenci değerlendirmesi tasarlama

B. Öğretimi Gerçekleştirme

- B1. Öğrenmeye Güdüleme ve Hedeften Haberdar Etme
- B2. Öğretmen-Öğrenci Etkileşimi
- B3. Öğretim Yöntem ve Tekniklerini Uygulama
- B4. Öğrenme Ortamını Düzenleme
- B5. Bilgiyi Anlamlandırma ve Üst Düzey Düşünme Becerilerini Geliştirme

B6. Öğrenmeyi değerlendirme

C. Yansıtıcı Düşünme

C1. Kişisel performansı değerlendirme

C2. Profesyonel gelişim planını oluşturma

D. İletişim ve İş birliği

D1. Olumlu iletişim kurma

D2. Meslektaşlardan gelen geri bildirimlere açık olma

D3. Meslektaşları ile İş birliği İçinde Çalışma

E. Mesleki Sorumluluklar

E1. Mesleki Gelişim

E2. Okul İçi Sorumluluklar

Yeterlik Ölçütleri	Değerlendirici	Yöntem ve Veri Toplama Aracı	Ürün	Birinci Dönem	İkinci Dönem
A1 ve B5	Farklı Branştan Öğretmen	Gözlem (Öncesi ve Sonrasında Görüşme)	Rapor	Ekim	Şubat
A1-A2-A3-A4-A5-A6-B1-B2-B3-B4-B5-B6-C1-D1-D2-D3-D4	Öğretmenin kendisi	Öz-değerlendirme	Rapor		
A1-A2-A3-A4-A5-A6-B1-B2-B3-B4-B5-B6	Başka bir sınıf öğretmeni	Gözlem (Öncesi ve Sonrasında Görüşme)	Rapor	Kasım	Mart
A1-A2-A3-A4-A5-A6-B1-B2-B3-B4-B5-B6	Yöneticiler (Müdür ve Müdür Yardımcısı)	Gözlem (Öncesi ve Sonrasında Görüşme)	Rapor		
D1-D2-D3-D4	Zümre Başkanı	Gözlem	Toplantı Notlar ve Rapor		
	Başka bir sınıf öğretmeni	Gözlem	Zümre toplantı saatleri ve ders dışı etkileşim		
A1-A2-A3-A4-A5-B1-B2-B3-B4-C1-D1-D2-D3-D4	Öğretmenin kendisi	Öz-değerlendirme	Rapor	Aralık	Nisan
A1-A2-A3-A4-A5-A6-B1-B2-B3-B4-B5-B6	Yöneticiler (Müdür ve Müdür Yardımcısı)	Gözlem (Öncesi ve Sonrasında Görüşme)	Rapor		
A1-A2-A3-A4-A5-A6-B1-B2-B3-B4-B5-B6	Başka bir sınıf öğretmeni	Gözlem (Öncesi ve Sonrasında Görüşme)	Rapor	Ocak	Mayıs
E1-E2	Zümre Başkanı	Gözlem	Rapor	-	Mayıs
	Müdür ya da Uzman	Bir yılda önerilen ve yürütülen eğitimlerin, projelerin, çalışmaların vb. incelenmesi.	Rapor	-	Mayıs
C2	Müdür	Performans Değerlendirme Görüşmesi (toplantısı)	Gelişim Planı*	Ocak	Mayıs

H. PART OF THE MODEL'S GUIDE

Öğretmen Mesleki Gelişimini Desteklemek için Katılımcı Bir Öğretmen Değerlendirme Modeli

I. Giriş

Bir eğitim sistemi, öğrenci performansı öğretmenleri tarafından iyileştirildiğinde ve her öğrencinin tüm potansiyelini geliştirildiğinde etkilidir. Diğer bir deyişle, öğretmenler eğitim sistemlerinin merkezinde yer alan en etkili güç olarak değerlendirilmekte ve öğretmenlerin yetenek ve niteliklerinin öğrencilerin öğrenmelerine önemli katkılar sağladığı düşünülmektedir (Darling-Hammond, 2006; Fullan, 2011; Wiliam, 2018). Öğretmenin etkililiğinin sadece sınıf ortamında geçerli olduğu düşünülemez. Öğretmenin etkisi sınıfın duvarlarını aşar ve öğretmenler de öğrencileri adım adım gelecekte onları bekleyen durumlara hazırlar. Öğretmenler, öğrencilerinin ihtiyaçları doğrultusunda öğretimi planlayarak, uygulayarak, değerlendirerek ve kararlar alarak öğrencilerinin çağdaş dünyanın karmaşık ve belirsiz koşullarına uyum sağlamalarına öncülük etmelidir. Eğitimin, ekonomiyi ve ulusları şekillendirmedeki önemli rolü nedeniyle en etkili öğretmenlere ve liderlere sahip olan okulların gelecekte de söz sahibi olacağını göstermektedir (Stronge, 2018).

Okulları iyileştirmek, öğrenmeyi iyileştirmek ve öğrenci başarısını arttırmak için öğretmenin öğrenmesi ve gelişmesi gereklidir. Öğretmenler öğretme zanaatında uzmanlık geliştirirlerse ve hangi öğretim stratejilerinin öğrencinin öğrenmesi üzerinde en olumlu etkiye sahip olduğu konusunda bilgi sahibi olurlarsa, öğrenci başarısı artacaktır (Netolicky, 2020). Bu amaçla öğretmeni bireysel olarak geliştirmek önemlidir. Eğitim alanında başarılı ülkeler, iyi öğretmenlere sahip oldukları için değil öğretmeni geliştirmek üzerine kurdukları yapıyı tüm öğretmenleri bireysel olarak geliştirmek üzerine kurdukları için başarılı oldular (Fullan, 2011).

Öğretmen kalitesi, öğrenci başarısını etkileyen en önemli faktör olarak tanımlandığından, öğretmen değerlendirme sistemleri aracılığıyla öğretmen kalitesini teşvik etmeye vurgu yapılmıştır (Looney, 2011; Papay, 2012). Elbette tüm öğrenciler ve sınıflar eşsizdir ve öğretmenlerin öğretim faaliyetlerinin tek bir etkili yolunun olduğundan söz etmek mümkün

değildir ancak bu, öğretmenlerin mevcut etkililiğinin arttırılamayacağı anlamına gelmemektedir. Öğretmenlerin etkililiğini arttırmak ve mesleki gelişimini desteklemek için de öğretmenlerin güçlü ve geliştirilmesi gereken yönlerinin doğru tespit edilmesine ihtiyaç vardır. Bu nedenle bir kurumda öğretmenin gelişimini sağlamak için çalışmalar yürütülmeden önce, etkili bir değerlendirme modeli ile öğretmenlerin hangi alanlarda ve ne tür eksiklikleri olduğunu belirlemek oldukça önemlidir. Genel olarak öğretmen değerlendirmesi, mesleki gelişim için yapıcı geri bildirim ve rehberlik sağlamak için hem okulda hem de sınıfta öğretmenlerin performansını gözden geçirmek için kullanılan sistematik bir süreç olarak tanımlanabilir. Eğitim politikalarını geliştirenler ve akademisyenler, öğretmen değerlendirme sistemlerini, öğrencilerin akademik performansını yükseltmek için öğretmenlerin kalitesini iyileştirmenin çok önemli bir parçası olarak tanımlamaktadır (Gordon, et al., 2006).

Öğretmenlik mesleğini geliştirmede uygulanan en büyük hatalardan biri genel ve evrensel olanın doğruluğuna inanmak ve bunu olduğu gibi uygulamaktır (Lewis & Hogan, 2016). Okullar öğretmeni, yöneticisi, öğrencisi ve okul kültürü ile bir bütün olarak düşünülmeli, mesleki gelişime odaklanan öğretmen değerlendirme yaklaşımları, okulun bağlamına ve hedeflerine uygun olarak tasarlanmalıdır (OECD, 2013a). Tüm okullar aynı sistemin bir parçası olmasına rağmen, içinde bulundukları farklı çevresel koşullar, öğrencilerin ve öğretmenlerin ihtiyaçları nedeniyle her okulun mesleki gelişim ihtiyaçları ve gelişim planları farklı olacaktır (TEDMEM, 2018). Okullar için öğretmen değerlendirmeye yönelik hazırlanmış modeli olduğu gibi alıp uygulamak her ne kadar kendi modelini oluşturmaktan daha kolay ve kullanışlı görünse de bu uygulamalar okulun ihtiyaçlarını karşılamamakta, beklenen etkiyi göstermemekte ve bu tarz modellerin kullanılmasından zamanla vazgeçilmektedir (Ofsted, 2018).

Okula özgü geliştirilen bir değerlendirme modelinde bu modelin uygulayıcılarının fikirlerini almak onların oluşturulan modele güven duymalarını sağlarken, modeli uygulamaya istekli hale getirir (Danielson & McGreal, 2000; Marzano and Toth, 2013). Bu öğretmen değerlendirme modeli oluşturulurken modelin her bir bileşeni için ve bu modelin nasıl uygulanması gerektiğine yönelik kurumunuzda çalışan öğretmenlerin, yöneticilerin ve uzmanların görüşleri alınmıştır. Elde edilen görüşler doğrultusunda model kapsamında değerlendirilecek öğretmen yeterlik alanları, alt alanlar, ölçütler, bu ölçütlerin kullanıldığı veri toplama araçları (gözlem formları, öz değerlendirme formlar vb.) oluşturulmuş, modelin amacına ve uygulanışında dikkat edilmesi gereken hususlara birlikte karar verilmiştir. Kuruma

özgü geliştirilmiş bu öğretmen değerlendirme modelinin her bir bileşeni ile ilgili ayrıntılı açıklama aşağıda yer almaktadır.

II. Öğretmen Değerlendirme Modelinin Amacı

Mesleki gelişimlerinin desteklendiği okul ortamında çalışan öğretmenlerin, daha az destek gören okullarda çalışan öğretmenlere göre çok daha hızlı geliştirdikleri bilinmektedir (Stronge, 2018). Öğretimin iyileştirilmesi ve öğretmenin mesleki gelişiminin sağlanması amacıyla yürütülen değerlendirme sistemlerinde, öğretmenlere etkili geribildirim vermeye ve onlara rehberlik etmeye odaklanılır (Kane vd., 2014; OECD, 2009b). Öğretmenin mesleki gelişimini temel alan modellerde yeterlik alanları, yönetici ve öğretmenler tarafından yapılan sınıf içi gözlemler, öz değerlendirme ve görüşmeler ile belirlenebilir. Değerlendirme öğretmenin mesleki gelişimini desteklemek amacıyla yapıldığında ve değerlendirmenin sonuçları öğretmenlerin uygulamalarını geliştirmek amacıyla kullanıldığında öğretmen, yapılan her türlü değerlendirmenin ve alınan kararın kendisini geliştirmek amacıyla olduğunu ve bu kararlar doğrultusunda bir yaptırım uygulanmayacağını farkında olarak kendi uygulamalarını içtenlikle sorgular, eksik olduğu yönlerini göstermekten çekinmez ve kendi öğretmenliği ile ilgili daha etkili kararlar alır (OECD, 2009b; OECD, 2013; TEDMEM, 2018). Bu değerlendirme modelinde de öğretmen niteliklerinin değerlendirilmesi ve değerlendirme sürecine bağlı olarak öğretmen mesleki gelişimi için sistematik destek sağlanması amaçlanmaktadır.

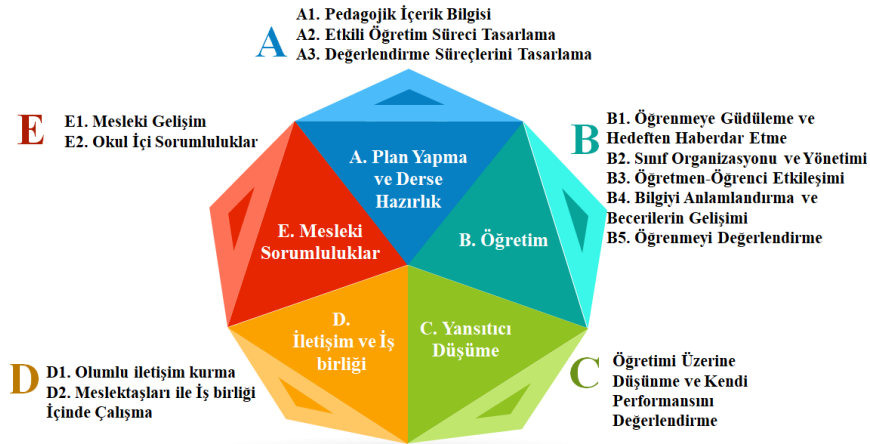
Modelin amacı doğrultusunda yapılan gözlem, görüşme ve öz değerlendirme sonuçları yılda iki kez (ocak ve mayıs aylarında) raporlanarak öğretmenlerle paylaşılacak ve öğretmenler ile yapılan “Performans Değerlendirme Görüşmesi” sonucunda her öğretmen için bir gelişim planı hazırlanacaktır. Bu gelişim planı, öğretmenlerin geliştirmesi gereken yeterlik alanları ve ihtiyaçları doğrultusunda belirlenen hedefler ve bu hedefleri gerçekleştirmek için yapılacaklara yönelik önerilerin yer aldığı bir plandır. Bu plan doğrultusunda öğretmenler belirlenen hedefleri gerçekleştirmek için bir dönem süresince eğitimlere katılabilir, meslektaşları ile iş birliği içinde çalışabilir, araştırmalar yapabilir, sınıf gözlemlerine katılabilirler. Bu öğretmen değerlendirme modelinin bileşenleri şekil 1 de yer almaktadır.



Şekil 1: Modelin Bileşenleri

III. Öğretmen Yeterlik Alanları, Yeterlik Alt Alanları ve Ölçütler

Bir öğretmen değerlendirme modeli öğretmen yeterlik alanları ve alt yeterlik alanları açısından hem spesifik hem de kapsamlı olmalıdır. Etkili öğretmenlik davranışlarını kapsayan ve bu davranışları yansıtan bir öğretmen değerlendirme modelinin ölçüt tabanlı olması ve bu ölçütlerin kapsamlı oluşturulması önemlidir. Bu değerlendirme modelinde yer alan ve aşağıda açıklanan yeterlik alanları, alt yeterlik alanları ve ölçütler kurumunuzdaki öğretmenlerin, yöneticilerin ve uzmanların görüşleri doğrultusunda oluşturulmuştur.



Şekil 2: Modelde yer alan yeterlik ve alt yeterlik alanları

IV. Öğretmen Değerlendirme Süreci

Bu değerlendirme modeli kapsamında öğretmenler okul müdürü, müdür yardımcıları, başka bir sınıf öğretmeni, farklı branştan bir öğretmen ve kendileri tarafından değerlendirilecektir. Değerlendirmeler yapılırken gözlem, görüşme ve öz değerlendirme formları kullanılacaktır. Öğretmenlerin ders planları ise ders planı inceleme formu yardımıyla değerlendirilecektir.

Eğitimler

Değerlendirmede bulunacak ve değerlendirilecek kişilere düzenli aralıklarla verilen bu eğitimlerde değerlendirme modelinin amacı, süreci, veri toplama araçlarının etkili kullanımı ve raporlama süreci hakkında uygulamalı eğitimler verilir. Gözlemin bu değerlendirme modelinin önemli bir parçası olduğu düşünüldüğünde, model uygulanmadan önce yapılacak eğitimlerde sınıf ortamında çekilmiş videolar, sınıf gözlemine yönelik yazılmış transkriptler, varsayımsal senaryolar gibi örnekler yardımıyla gözlem formların kullanımı uygulamalı olarak gerçekleştirilir ve değerlendiricilerin bir ders gözlemine aynı bakış açısıyla bakabilmesine rehberlik edilir. Eğitimler her yıl belli aralıklarla düzenli olarak tekrarlanır ve gelen öneriler doğrultusunda eğitimlerin içeriği güncellenir.

İş ve zaman çizelgesi

Öğretmenlerin okul açılışının ilk aylarında yaşadığı oryantasyon süreci göz önüne alındığında, modeldeki gözlem süreçleri her yarıyılın ilk ayında yapılmaz. Değerlendirmelerin hangi ayda, kim tarafından, hangi yeterlik alanına yönelik ve ne tür formlar kullanılarak yapılacağına ilişkin bilgiler aşağıda yer almaktadır.



Veri Toplama Araçları

Geliştirilen veri toplama araçları gözlem, görüşme formları, öz değerlendirme formu ve ders planı inceleme formudur. Tüm araçlar ve araçlarda yer alan değerlendirme ölçütleri bu okulda görev yapmakta olan sınıf öğretmenlerinin, uzmanların ve yöneticilerin görüş ve katkılarıyla oluşturulmuştur. Kullanılacak değerlendirme araçları Ek-F’de yer almaktadır. Bu veri toplama araçlarının ne amaçla ve nasıl kullanılacağı aşağıda açıklanmaktadır.



Şekil 3: Modelde kullanılan veri toplama araçları

Veri Toplama Süreci

Modelde kullanılan her veri toplama aracının ön sayfasında, nasıl kullanılacağı, dikkat edilmesi gereken önemli noktalar hakkında bilgi ve prosedürler ayrıntılı olarak verilmiştir. Bu model kapsamında mesleki tecrübesinden bağımsız olarak her öğretmen için aynı veri toplama aracı, aynı prosedürler doğrultusunda uygulanır. Bir öğretmen, güvenilir bilgi elde etmek ve değerlendirmelerin önyargıdan bağımsız olmasını sağlamak için birden çok değerlendirici tarafından birçok kez değerlendirilir.



Tutarlılık

Ders gözlem sürecinde belirlenen yeterlilik alanındaki ölçütlere uygun olarak öğretmen hakkında doğru kararlar verebilmek için, öğretmenin davranışının sürekli olup olmadığı kontrol edilmelidir. Diğer bir deyişle, öğretmenin ölçütlerde belirtilen davranışları her seferinde farklı seviyelerde gerçekleştirmesi durumunda davranışın istikrarı

hakkında konuşmak doğru olmaz. Ancak, değerlendirmeler tutarlı ise, yani, o ölçüt için yapılan her değerlendirmede öğretmenin seviyesi benzer ise, bu ölçüt ile ilgili yorumlar dikkate alınmalıdır. Bu durumun çözümü için, değerlendiricilerin formları etkin bir şekilde kullanmasını sağlamak amacıyla videolara, sınıf gözlemine yönelik videolara, yazılmış transkriptlere, varsayımsal senaryolara dayalı değerlendirme yapma şansına sahip olacaklardır. Ayrıca, özellikle gözlemler yoluyla elde edilen sonuçların faydalı ve geçerli olabilmesi için öğretmenlerin sınıf performansı, farklı gözlemciler tarafından benzer güvenilirliklerin elde edilmesi için en az dört kez gözlemlenecektir.

Etkili bir öğretmenlik, sınıfta gerçekleştirilen etkinliğin türüne, konu alanına, öğrencinin önceki öğrenmelerine ve karakterlerine bağlıdır (Jones, Jenkin & Lord, 2006). Bu nedenle, bir sınıf öğretmenin etkinliğini gözlemlerken, öğretmenin farklı dersler, farklı konu ve farklı ders bölümlerindeki performansı (yeni bir konu öğretilmesi, bir konuyu pekiştirme, değerlendirme vb.), birden fazla kişi tarafından değerlendirilmelidir. Gözlemlenecek ders ve sınıftaki öğrencilerin özellikleri ile ilgili bilgiler, gözlemden önce yapılan görüşmeler ile toplanır ve ders bu bilgiler doğrultusunda izlenir. Bu modelde kullanılan formlar, öğretimin değerlendirilmesinde belirtilen bu etkili yolun kullanılmasını sağlayacak şekilde oluşturulmuştur.



Objektiflik

Aynı seviyelerde öğretim yapan gözlemci ve gözlenen kişiler arasındaki kişisel ilişkilerin değerlendirme sürecini etkilememesi ve değerlendirmenin tarafsızlığının sağlanması amacıyla ders gözlemleri farklı seviyelerde görev yapan sınıf öğretmenleri tarafından gerçekleştirilir. Değerlendirilecek öğretmenin bir sınıf öğretmeni olması, sınıf öğretmenlerinin her yıl farklı seviyelerde ders vermesi

ve dört seviyenin tümüne hâkim olması bu durumun gerçekleştirilmesine olanak sağlamaktadır. Ayrıca gözlemin iki farklı gözlemci tarafından aynı anda yapılması sağlanarak değerlendirmelerin tarafsızlığını sağlar. Aynı dersi gözlemleyen farklı iki kişi arasındaki tutarlılığın sağlanmasında değerlendirme öncesinde alınan eğitimlerin etkili olacağı düşünülmektedir.

V. Değerlendirme Sonuçlarının Paylaşımı



Açıklık

Her gözlem ve görüşmeden sonra bu model kapsamında öğretmenlere aynı gün geri bildirim verilir ve bu geri bildirim verme süreci sadece değerlendirici ile öğretmen arasında yapılır. Tüm değerlendirme sonuçlarının öğretmene yol göstermesi ve faydalı olması için, değerlendirmeler ölçeklerde açıkça belirtilen ölçütler yardımıyla ve bu ölçütlerin gerçekleştirilme düzeyi belirtilerek paylaşılır.



Gizlilik

Tüm değerlendirmelerden elde edilen görüşlerden sonra bir rapor hazırlanır ve bu raporlar her öğretim yılının ocak ve mayıs ayı içerisinde okul müdürü ile yapılan yüz yüze görüşmelerle öğretmenle paylaşılır. Bu paylaşımın yapıldığı “Performans Değerlendirme Toplantısı’nda öğretmenler her zaman değerlendirme sonuçları hakkındaki görüşlerini ifade etme hakkına sahiptir. Bu toplantının bir amacı, öğretmenlerin güçlü yönlerini takdir etmek ve diğer tüm öğretmenlerin bu yönlerden faydalanmasını sağlamak için iş birliğine dayalı bir çalışma planı oluşturmaktır. Bu toplantının ikinci amacı ise öğretmenin belirlenen eksiklikleri ile ilgili bir dönem boyunca neler yapılabileceğini planlamak ve böylece öğretmenin mesleki gelişimini desteklemektir. Bu iki amaç doğrultusunda öğretmen ile birlikte bir mesleki gelişim planı hazırlanır. Yapılan değerlendirmelerin ve hazırlanan gelişim planının içeriği veya mesleki gelişim planlaması süreci tamamen gizli tutulur, bu içeriğe yalnızca değerlendirilen öğretmen erişebilir.

I. EXAMPLES OF DATA COLLECTION TOOLS OF THE MODEL

İletişim ve İş birliği Değerlendirme Formu

Değerlendirici: Zümre Başkanı, Sınıf Öğretmeni

Yönerge: Bu değerlendirme formu ile belirtilen ölçütler doğrultusunda öğretmenin iletişim ve iş birliğine yönelik durumunun değerlendirilmesi amaçlanmaktadır. Buna göre, ölçütte belirtilen durumun ilgili öğretmen tarafından gerçekleştirilme sıklığını düşünerek “Hiç”, “Nadiren”, “Sık sık” ve “Tamamen” kategorilerinden birini işaretleyiniz.

	Ölçütler	Hiç	Nadiren	Sık sık	Tamamen
Olumlu iletişim kurma	Karşısındakini dikkatle dinlemektedir.				
	Dinlediğini anlamaya yönelik sorular sormaktadır.				
	Empati kurduğunu gösteren ifadeler kullanmaktadır.				
	Yargılayıcı ya da öğüt verici ifadeler kullanmamaktadır.				
	Eleştirilerinde yapıcı bir dil kullanmaktadır.				
	Yorum yaparken olumlu ifadeler kullanmaktadır.				
	Olayları farklı bakış açılarıyla değerlendirmektedir.				
İş birliği içinde Çalışma	Etkili olduğunu düşündüğü ders planı/öğretim süreci/ materyali/ çalışmayı paylaşmaktadır.				
	Meslektaşından öğrenmeye açıktır.				
	Kendisine sunulan fikirlere, bildirimlere açıktır.				
	Meslektaşlarından yardım almaya isteklidir.				
	Meslektaşlarına ihtiyaç duyduklarında yardım etmektedir.				
	Grup çalışmalarında aktif, etkin görev almaktadır.				
	Grup çalışmalarında verilen görevi zamanında tamamlamaktadır.				
	Destek birimler (PGÖDEM, PDRM) ile uyumlu ve iş birliği içinde çalışmaktadır.				

Mesleki Sorumluluk Değerlendirme Formu

Değerlendirici: Müdür, Müdür Yardımcısı

Yönerge: Bu değerlendirme formu ile belirtilen ölçütler doğrultusunda öğretmenin mesleki sorumluluğunun değerlendirilmesi amaçlanmaktadır. Buna göre, ölçütte belirtilen durumun ilgili öğretmen tarafından gerçekleştirilme sıklığını düşünerek “Hiç”, “Nadiren”, “Sık sık” ve “Tamamen” kategorilerinden birini işaretleyiniz.

	Ölçütler	Hiç	Nadiren	Sık sık	Tamamen
Mesleki Gelişim	Yeterliğini geliştirmek için seminer, kongre sempozyum gibi eğitimlere katılmaktadır.				
	Alanı/mesleği ile ilgili yayınları takip etmektedir.				
	Alanı/mesleği ile ilgili gündemi ve gelişmeleri takip etmektedir.				
	Katıldığı eğitimler, takip ettiği yayınlar, uygulamalar vb.’den edindiği bilgileri çalışmalarına yansıtmakta ya da öneri getirmektedir.				
	Gelişme planındaki dönütler doğrultusunda çalışmalarını düzenlemektedir.				
Okul İçi Sorumluluklar	Okul kurallarına uymaktadır.				
	Toplantılara katılmakta ve alınan kararları uygulamaktadır.				
	Görev ve sorumlulukları (yıllık plan-günlük plan hazırlama, ödev hazırlama, etkinlik düzenleme, nöbet tutma, tören hazırlıkları vb.) yerine getirmektedir.				
	Kendisine verilen görevlerde yaptığı işler niteliklidir.				
	Kendisine verilen işleri zamanında yerine getirmektedir.				
	Öğrencilerin sosyal gelişimine yönelik kanıtlar toplayarak gelişimlerini takip etmektedir.				
	Öğrencilerin akademik gelişimine yönelik kanıtlar toplayarak gelişimlerini takip etmektedir.				
	Kaynakları kullanırken tutumlu davranmaktadır.				
	Çevresini temiz ve düzenli tutmaktadır.				

J. APPROVAL OF THE METU HUMAN SUBJECTS ETHICS COMMITTEE

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER



ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

Sayı: 28620816 / 61
DÜMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T: +90 312 210 22 91
F: +90 312 210 79 59
Konu: Değerlendirme Sonucu
www.usam.metu.edu.tr

02 OCAK 2018

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi: İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Doç.Dr. Yeşim ÇAPA-AYDIN;

- Danışmanlığınızı yaptığınız Emel Güneş SAVUL'un "**Öğretmenlerin Değerlendirilmesi ve Öğretmen Mesleki Gelişim için Bir Model: Örnek Olay İncelemesi**" başlıklı araştırması İnsan Araştırmaları Etik Kurulu tarafından uygun görülerek gerekli onay **2017-EGT-200** protokol numarası ile **02.01.2018-30.12.2019** tarihleri arasında geçerli olmak üzere verilmiştir.

K. APPROVAL OF THE PRIVATE INSTITUTION (CASE)



Sayı : B.08.4.MEM.2.06.51.04-900- 226 27/02/2018
Konu : Emel GÜNEŞ SAVUL'un
Doktora Çalışması

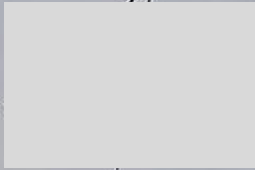
Genel Müdürlüğe

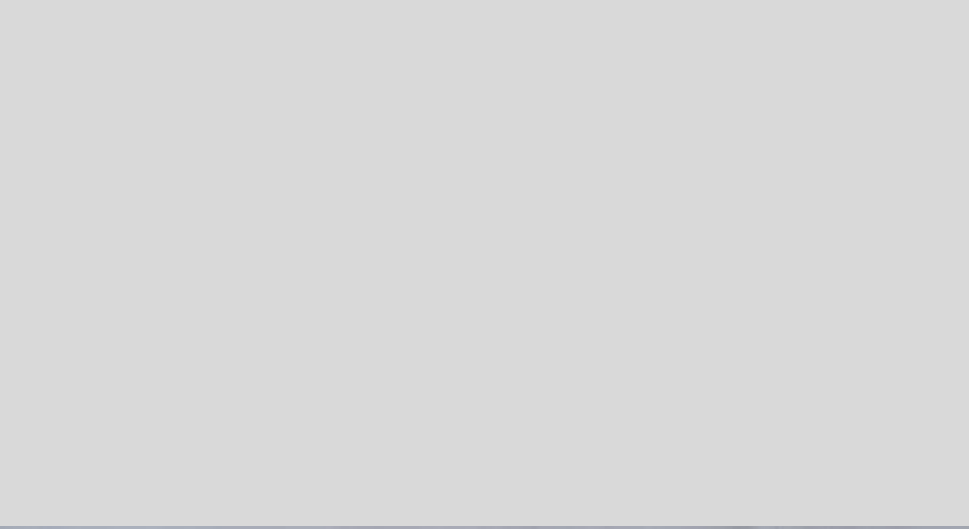
İlgi : Program Geliştirme Uzmanı Emel GÜNEŞ SAVUL'un 23/02/2018 tarihli dilekçesi.

Kısmımız Program Geliştirme Uzmanı Emel GÜNEŞ SAVUL, Orta Doğu Teknik Üniversitesi, Eğitim Bilimleri Bölümü, Eğitim Programları ve Öğretimi Programı'nda doktora öğrenimine devam etmektedir. Adı geçen uzmanımız ilgide kayıtlı dilekçesinde, "Öğretmenlerin Değerlendirilmesi ve Öğretmen Mesleki Gelişimi İçin Bir Model: Örnek Olay İncelemesi" adlı tez çalışması kapsamında, İlkokul kısmımız yöneticileri, sınıf öğretmenleri ve PGÖDEM uzmanlarından random olarak seçilecek bir grup ile ekte sunulan görüşme formlarını uygulamak istemektedir.

Onayınıza arz ederim.

Eki : Dilekçe ve Görüşme Formları (7 Sayfa)







L. CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Em

Nationality: Turkish (TC)

Date and Place of Birth:

Marital Status: Married

Phone:

email:

EDUCATION

Degree	Institution	Year of Graduation
MS	Hacettepe University Curriculum and Instruction	2009
BS	Hacettepe University Science Education	2006
High School	Ayşe Atıl High School, Adana	2001

WORK EXPERIENCE

Year	Place	Enrollment
2014 Agust	TED Ankara College	Curriculum and Development Specialist
2008 Agust	Gazi University Foundation Private Schools	Curriculum and Development Specialist

FOREIGN LANGUAGES

Advanced English

PUBLICATIONS

1. Kalaycı, N. And Güneş, E. “Determination the Effects of Hidden Messages in Official Curriculum on Students’ Perceptions of the Freedom Value Freedom or Safety?”, Türk Eğitim Bilimleri Dergisi, 12(1), 17-34 (2014)

HOBBIES

Dance, Scuba, Fine Arts, Photography, Music, Tracking

M. TURKISH SUMMARY / TRKE ZET

GİRİŞ

ğretmenler eğitim sistemlerinin merkezinde yer almakta ve ğrenci başarısı için en etkili güç olarak değerlendirilmektedir (Fullan, 2011; OECD, 2018; Stronge, 2006; Wiliam, 2018). Etkili ğretim tarih boyunca farklı şekillerde tanımlanmıştır. ğretmenin bilginin bekçisi olarak görüldüğü ve etkili ğretmenin de ğrencinin kontrollü bir şekilde bilgiye ulaşmasını sağladığı görüşü yerini yeni tanımlara bırakmıştır. Günümüzde hem beynin nasıl ğrendiğine ilişkin nöroloji alınındaki araştırmalar hem de teknoloji alanındaki gelişmeler, ğrenmede ğretmen-ğrenci etkileşiminin önemli olduğunu göstermektedir. Dahası bilgiye ulaşabilen, kendi çabası ve tecrübesiyle ğrenebilen, eleştirel düşünme ve problem çözme gibi üst düzey düşünme becerilerine sahip ğrenciler yetiştiren ğretmen etkili bir ğretmen olarak tanımlanabilir. Etkili ğretim ve ğretmen yıllar içinde birçok farklı tanımlı beraberinde getirmiş olsa da etkili bir ğretmenin sahip olması gereken niteliklerin ğrencilerin ğrenmelerine ve başarılarına olan katkılarla birlikte belirlendiği söylenebilir (CDE, 2015; Clotfelter ve diğerleri, 2007; Darling-Hammond, 2006; Hanushek, 1992; Kim ve Sun, 2021). Buradaki “katkı” sözcüğü önemlidir, çünkü ğretmen ğrenciye sadece mevcut ğrenme sürecinde değil, aynı zamanda gelecekteki ğrenmelerine de katkı getirir. Tucker ve Stronge'un (2005) belirttiği gibi, ğrencilere bir yıl boyunca yüksek performans düzeyine sahip ğretmenler ile ğretim gördüğünde, ilerleyen yıllarda bu etkili ğretimden yararlanmaya devam ederler, tam tersine düşük performans gösteren bir ğretmen tarafından sağlanan ğretimin istenmeyen etkilerini de ilerdeki ğrenmelerine taşırlar ve bu durum ğrencinin sonraki yıllarda ğrenmesini olumsuz etkiler. Dahası ğretmen etkililiğinin sadece sınıf ortamında geçerli olduğu düşünülemez. ğretmenin etkisi sınıfın duvarlarını aşar ve ğretmenler de ğrencileri adım adım gelecekte onları bekleyen durumlara hazırlar. Eğitimin, ekonomiyi ve ulusları şekillendirmedeki önemli rolü nedeniyle en etkili

öğretmenlere ve liderlere sahip olan okulların gelecekte de söz sahibi olacağını göstermektedir (Stronge, 2018). Öğretmenler aynı zamanda günümüzün rekabetçi dünyasının vazgeçilmez bir unsuru olan insan gücünün sağlanmasında da önemli bir rol oynamaktadır ve toplumlar öğretmenlik mesleğine yoğun bir şekilde yatırım yapmaktadır (OECD, 2019; Wiliam, 2018).

Öğretim sürecinin etkililiğinin öğretmenin etkililiği ile doğrudan ilişkili olması nedeniyle öğretmenin gelişimini ve ilerlemesini sağlayarak daha başarılı öğrenciler yetiştirmek en umut verici ve temel stratejilerden biridir. Öğretmenin yeterliklerini ve dolayısıyla öğretimin kalitesini iyileştirme fikri, politika yapıcılar için önemli bir odak noktasıdır ve öğretmen iyileştirilmesi gereken birincil araç olarak kabul edilir (CDE, 2015; Darling-Hammond, 2010; Darling-Hammond, 2014; Gordon ve diğerleri, 2006; Stronge, 2006; Şahan, 2011). Öğretmenin gelişimini desteklemenin başlangıç noktası değerlendirmedir. Değerlendirilmeyen bir şey değiştirilemez veya geliştirilemez bu nedenle de öğretmenin yeterliklerinin iyileştirilmesi de öğretmenin değerlendirilmesi ile mümkündür (Looney, 2011; Papay, 2012). Genel olarak öğretmen değerlendirmesi, mesleki gelişim için yapıcı geri bildirim ve rehberlik sağlamak amacıyla hem okulda hem de sınıfta öğretmenlerin performansını gözden geçirmek için kullanılan sistematik bir süreç olarak tanımlanabilir (Ornstein, 1985; Danielson, 2011; Darling-Hammond, 2010).

Değişen eğitim politikaları ve reformlar okullarda uygulanan öğretmen değerlendirme sistemlerinin güncellenmesine veya yeniden yapılandırılmasına neden olmuştur (Dee vd., 2021; Donaldson ve Woulfin, 2018; Garrett ve Steinberg, 2015). Ancak yapılan reformların veya uygulamaların eğitim sistemini geliştirmede ve daha etkili öğretmen değerlendirme modelleri ortaya koymada beklendiği kadar etkili olmadığı söylenebilir. Özellikle Amerika Birleşik Devletleri'nde reformların etkisinin beklenen düzeyde olmamasının en önemli nedenlerinden biri, ölçütleri temel alarak yapılan değerlendirmelerden ziyade katma değer puanları (value added scores) gibi normatif ölçümlere daha fazla ağırlık verilmesidir (Kraft ve Gilmour, 2017). Bugün eyaletler Katma Değer Modellerinin (Vakue Added Models) ağır yükünü hafifletmenin yollarını aramaya başlamıştır. Birçok eyalet, öğrenci başarısı ile öğretmen etkinliği

arasındaki ilişkileri ölçerek öğretmen değerlendirmesine daha fazla alternatif sunmaktadır ve eyalet öğretmen değerlendirme planları daha biçimlendirici öğretmen geribildirimini içermektedir (Close vd., 2020). Değerlendirmeyi yapan kişilerin eğitim eksikliği de bu reformların etkisini azaltan faktörler arasındadır.

Bu zamana kadar öğretmeni değerlendirmenin önemini kabul ederek, yapılan reformların ve geliştirilen modellerin belki de en önemli katkısı, etkili bir öğretmen değerlendirme modelinde olması gereken bazı temel bileşenlerin belirlenmesine yardımcı olmasıdır. Tek ve etkili bir öğretmen değerlendirme modelinden bahsetmek mümkün olmasa da ayrıntılı standartlar veya ölçütler yardımıyla öğretmen yeterliklerinin açıkça tanımlayarak, yıl boyunca ölçütlere dayalı gözlemler yaparak ve kanıt toplamak için birden fazla ölçüm kullanarak güvenilir öğretmen değerlendirme sistemleri geliştirilebilir (Darling-Hammond ve diğerleri, 2012; Weisberg ve diğerleri, 2009, Putman ve diğerleri, 2018). Ayrıca başarılı ve güçlü öğretmen değerlendirme modellerinin amacının, öğretmenin iyi performans gösterdiği alanları ve öğretmenin geliştirmesi gereken alanları belirlemek için öğretmenin mesleki gelişimi ile bağlantılı olduğunu söylemek de mümkündür (Putman vd., 2018). Genel olarak, öğretmenin gelişimini temel alan, zamanında biçimlendirici geri bildirim sağlayan, öğretmenlerin değerlendirme sürecine katılımını sağlayan, birden fazla kaynaktan kanıt toplayarak veri toplama işlemlerini yürüten değerlendirme süreci daha etkili ve güçlüdür (Danielson, 2013; Darling-Hammond, 2012; Derrington ve Brandon, 2019; Marzano, 2012; OECD, 2013a).

1.1 Çalışmanın Amacı

Bu araştırmanın amacı, öğretmenlerin mesleki gelişimini desteklemek için katılımcı bir öğretmen değerlendirme modeli oluşturmaktır. Araştırmanın ilk bölümü, sınıf öğretmenleri, yöneticiler ve uzmanların özel okulda mevcut öğretmen değerlendirme ve mesleki gelişim uygulamalarına ilişkin algılarının derinlemesine bir analizini yapmayı ve katılımcıların önerilerini derinlemesine incelemeyi amaçlamıştır. Araştırma soruları şunlardı:

1. Bu özel okulda görev yapmakta olan sınıf öğretmenleri için ne tür öğretmen değerlendirme süreçleri yürütülmektedir?
2. Öğretmen değerlendirme süreci bu özel okulda görev yapmakta olan sınıf öğretmenleri, yöneticiler ve uzmanlar tarafından güçlü ve zayıf yönler açısından nasıl algılanmaktadır?
3. Etkili bir öğretmen değerlendirme modeli geliştirmek için bu okulun sınıf öğretmenleri, yöneticiler ve uzmanlarının önerileri nelerdir?
4. Bu özel okulda ne tür mesleki gelişim süreçleri yürütülüyor ve nelere ihtiyaç var?

Bu araştırmanın ikinci ve üçüncü bölümünde, veri toplama araçlarıyla öğretmenlerin mesleki gelişimlerini desteklemek için katılımcı bir öğretmen değerlendirme modeli geliştirmek amaçlanmıştır. Araştırma soruları şunlardır:

5. Oluşturulan taslak öğretmen değerlendirme modeline ilişkin katılımcıların görüşleri nelerdir?
6. Oluşturulan öğretmen değerlendirme modelinin kılavuzuna ilişkin katılımcıların görüşleri nelerdir?
7. Öğretmen değerlendirme modelinde yer alan veri toplama araçlarının geçerlilik kanıtları nelerdir?

1.2 Çalışmanın Önemi

Tüm öğrencilerin yüksek standartlarda öğrenebileceğine inanılmasına rağmen, okullarda neden hala yetersiz öğretimin uygulandığını yanıtlamakta çok az ilerleme kaydedilmiştir. Daha da önemlisi, okullardaki bu yetersiz öğretimden sorumlu olanları iyileştirmek için öğretimi doğru bir şekilde değerlendirecek bir değerlendirme sistemi hala tam olarak geliştirilmemiştir (Grissom & Youngs, 2015; Kim & Sun, 2021; Weisberg ve diğerleri, 2009). Yetersiz kanıtlara dayalı, geçerli olmayan araçlarla verilerin toplandığı, öğretmenlere faydası olmayan geribildirimler sunan, çoğunlukla uzmanlıktan yoksun, yeterince eğitilmiş olmayan değerlendiriciler tarafından yürütülen ve alınan kararların dikkate alınmadığı değerlendirmeler kullanılmaya devam etmektedir (Jerald, 2009). Kısacası öğretmenin öğrenci başarısı üzerindeki gücünün

niteliğini belirlemek amacıyla onu doğru değerlendirmede ya da niteliğini geliştirmek amacıyla desteklemede kullanılan etkili sistemler geliştirildiğini söylemek pek mümkün değildir.

Öğretmenlerin değerlendirilmesi hemen hemen tüm ülkelerde saygın ve önemli bir konu olmasına rağmen belirtilen nedenlerden dolayı, mevcut öğretmen değerlendirme sistemleri öğretmen gelişimine çok az yardımcı olmakta ve öğretmenler hakkında karar verme süreçlerini zar zor desteklemektedir (Darling-Hammond ve diğerleri, 2012; NCTQ, 2011; NEA, 2012). Öğretmenler ayrıca net ölçütler belirlenmeden ve yöneticiler tarafından tek yönlü iletişim kullanılarak yapılan değerlendirmeler nedeniyle verilen dönütün kendine özgü ve dikkat çekici olmadığı gerekçesiyle değerlendirilmek istememektedirler (Danielson ve McGreal, 2000). Sonuç olarak öğretmenler, değerlendirmeleri kendi gelişimini desteklemek için işlevsel olmaktan çok teknik olarak algılamaktadırlar (McClure, 2008). Oysaki öğretmenler, mesleki gelişimlerini destekleyen değerlendirme modellerine açıktır (Campbell, 2014; Ramage, 2012; Passe, 2015). Araştırmalar, değerlendirme modellerinin gerekli becerilere sahip kişiler tarafından doğru bir şekilde uygulanması, öğretmenlere değerlendirme süreci konusunda eğitim verilmesi ve değerlendirmelerin öğretmenlerin değerlendiricilere güvendiği ve meslektaş dayanışması hissettiği bir ortamda yapılması durumunda öğretmenlerin güven oluşturabileceklerini ve gelişimleri için fırsatlar bulabileceklerini göstermektedir (Campbell, 2014; Rucinski ve Diersing, 2014; Walker, 2014). Dahası değerlendirme sistemlerinin dikkatli bir şekilde tasarlanıp uygulanırsa, öğretmen değerlendirme sistemleri okulun gelişimini teşvik etmede istenen sonuçlara ulaşabilir (Grissom & Youngs, 2015).

Problemi ve çözümü oldukça belli olan bu duruma rağmen öğretmen değerlendirme ile ilgili yapılan reformların emekleme aşamasında olduğu ve olgunluğa ulaşmak için birçok değişiklik ve yinelemeden geçmesi gerektiği ortadadır. Başka bir ifadeyle eğitim sistemlerinin etkili öğretmen değerlendirme sistemlerine ve modellerine ihtiyacı vardır (Marzano ve Toth, 2013). Etkili bir şekilde öğretebilen öğretmenlerin yetiştirilmesi için öğretmenlerin becerilerini doğru bir şekilde ölçmesi gerekir. Bu amaçla öğretmen performansını artırmayı sağlayan, standart temelli ve açık ölçütlere

dayalı deęerlendirmeler ieren ğretmen deęerlendirme modellerinde bir dnüşüme ihtiya vardır. Bu dnüşüm iin ğretmenlerin yetenek ve kapasitelerinin doęru bir şekilde anlaşılması ve deęerlendirilmesi gerekmektedir (Carbaugh vd., 2020; Darling-Hammond, 2010; Marzano, 2017). Bir dizi kontrol listesi kullanarak ğretmenlerin performanslarının farklı düzeylerini belirlemek yerine, okulların ok iyiye iyiden ve iyiye zayıftan ayırmasına ihtiyaı vardır (Danielson ve McGreal, 2000; Kraft ve Gilmour, 2017; Weisberg ve dięerleri, 2009).

Öğretmen deęerlendirme modellerinin hala istenilen düzeyde etkili olmamasının nedenlerinden biri deęişen eęitim politikaları kapsamında modeller hazırlanması; başka bir deyişle, modellerin harici bir kaynaktan önerilerek uygulanmasıdır (Kim & Sun, 2021). Eęitim politikalarının ve politikalar doęrultusunda harici olarak tasarlanan modellerin ye da uygulamaların etkililięini incelemek iin tüm aktrlerin yani uygulayıcıların deęişim ve katılım konusundaki isteklilięini incelemek gerekir (Viennet ve Pont, 2017). Bu nedenle, deęerlendirmeye ve etkisine ğretmenlerin bakış aısından bakılmalıdır (Derrington ve Brandon, 2019). Öğretmen ve yöneticiler gibi paydaşların deęerlendirme sistemlerinin tasarlanması, geliştirilmesi ve uygulanmasına katılımı deęerlendirme sisteminin uygulanabilirlięini artırır (Danielson & McGreal, 2000; Howard & Gullickson, 2009). Gelişimlerinde söz sahibi olan ğretmenler, uygulanan deęerlendirme sisteminin tamamen mesleki gelişimleri iin olduęuna ve sonuçların sadece kendi çıkarları iin kullanıldığına inanmaktadır (Attard, 2016; Brandenburg ve dięerleri, 2017).

Öğretmenlik mesleęini geliştirmede uygulanan en büyük hatalardan biri de genel ve evrensel olanın doęruluęuna inanmak ve bunu olduęu gibi uygulamaktır (Lewis & Hogan, 2016). Okullar ğretmeni, yöneticisi, ğrencisi ve okul kltürü ile bir bütün olarak düşünölmeli, mesleki gelişime odaklanan ğretmen deęerlendirme yaklaşımları, okulun bağlamına ve hedeflerine uygun olarak tasarlanmalıdır (OECD, 2013). Tüm okullar aynı sistemin bir parası olmasına rağmen, iinde bulundukları farklı çevresel koşullar, ğrencilerin ve ğretmenlerin ihtiyaları nedeniyle her okulun mesleki gelişim ihtiyaları ve gelişim planları farklı olacaktır (TEDMEM, 2018). Okullar iin ğretmen deęerlendirmeye yönelik hazırlanmış harici bir modeli olduęu

gibi alıp uygulamak her ne kadar kendi modelini oluşturmaktan daha kolay ve kullanışlı görünse de bu uygulamalar okulun ihtiyaçlarını karşılamamakta, beklenen etkiyi göstermemekte ve bu tarz modellerin kullanılmasından zamanla vazgeçilmektedir (Ofsted, 2018). Hem ilköğretim hem de ortaöğretim düzeylerinde, öğrencilerin, öğretmenlerin ve okul liderlerinin değerlendirilmesinde daha fazla yerleşme ve okul özerkliğine yönelik bir eğilim vardır. Etkili öğretmen değerlendirmesi, aynı okuldan görev yapan öğretmen ve yöneticilerin iş birliği içinde çalıştığı, sorunları birlikte çözmek için çaba gösterdiği, öğretmenin birbirinden öğrenmelerine imkan verildiği ve öğretmenin güven duyduğu bir okul ortamında yürütülebilir (Desimone, 2009). Öğretmene sınıf içi uygulamalara daha yakından bakma fırsatı sağlayan ve öğretmenler arasında iletişim ve iş birliği için fırsatlar sunan okul temelli bir değerlendirme, gelişim odaklı bir öğretmen değerlendirmesi oluşturmaya daha elverişlidir (Chen ve diğerleri, 2021). Okul temelli değerlendirmeler öğretmen gelişimi için elverişli bir ortam yaratır çünkü bu ortamda öğretmenler doğal öğrenendir ve sürece güven duyarak uygulamaya istekli hale gelebilir (Kurum ve Cinkir, 2019; Wong ve Li, 2010). Bu araştırma kapsamında, okuldaki ilgili tüm paydaşların katılımıyla ve mesleki gelişimi desteklemeyi amaçlayan, bir okulun yapısına ve kültürüne özel olarak geliştirilen ve önerilen öğretmen değerlendirme modelinin belirtilen sorunların çözümünde etkili olması beklenmektedir.

YÖNTEM

2.1 Çalışma Deseni

Nitel araştırma genellikle sorunu veya konuyu araştırmak için ayrıntılı ve karmaşık bir anlayışa ihtiyaç duyulduğunda kullanılır. Bu sorunu ayrıntılı bir anlayışla keşfederken, bireylerin veya bir grup insanın sosyal soruna yüklediği anlamı ele almak esastır (Marshall ve Rossman, 2015). Nitel araştırmalarda, çoklu bakış açılarını dahil etmek, araştırmaya dahil olan faktörleri belirlemek ve problemin veya konunun karmaşık bir resmini geliştirmek esastır (Creswell ve Poth, 2018). Bu çalışmanın nitel olmasının önemli bir nedeni, araştırmanın yapıldığı kurumda öğretmen değerlendirmesi ile ilgili mevcut uygulamaları ve bu uygulamaların güçlü ve zayıf yönlerini ayrıntılı olarak anlamayı amaçlamasıdır. Ayrıca, kurumlarında uygulanmak üzere tasarlanan etkili bir

öğretmen değerlendirme modeli için katılımcılardan öneriler alınması da hedeflenmiştir. Nitel araştırma seçilmesinin bir diğer nedeni de araştırma kapsamında katılımcı bir öğretmen değerlendirme modeli geliştirilmek istendiğinden, katılımcıların bireysel ve özgün görüşlerinin birçok kez alınmasına duyulan ihtiyaçtır. Ayrıca, katılımcı bir öğretmen değerlendirme modeli geliştirmek için farklı katılımcıların, çoklu bakış açılarının ve katılımcılar arasındaki etkileşimin önemli olduğu gerekçesiyle nitel araştırma yöntemi tercih edilmiştir. Bu çalışmada nitel yöntemlerle öğretmenlerin, yöneticilerin ve uzmanların mevcut değerlendirme süreçleriyle ilgili tecrübeleri ve farklı bakış açılarıyla birlikte ihtiyaç duyulan modele yönelik görüş ve önerileri de derinlemesine incelenmiştir.

Örnek olay incelemesi, araştırmacının gerçek hayattaki güncel bir durumu veya durumları, birden fazla bilgi kaynağı içeren ayrıntılı, derinlemesine bir veri toplama süreci aracılığıyla araştırdığı nitel yaklaşımlardan biridir (Creswell ve Poth, 2018). Araştırmacılar genellikle amaçlı örneklem yaparak bir bireyi, küçük bir grubu, bir kurumu veya bir organizasyonu vaka olarak seçebilir. Seçilen bu örneklem gereği vaka çalışmalarında asıl amaç genelleme değil, bu duruma ilişkin yeterli bir betimleme, yorum ve açıklama geliştirmektir (Creswell ve Poth, 2018; Marshall ve Rossman, 2015).

Öğretmenler, yöneticiler, örgüt kültürleri, çalışanlar arasındaki ilişkiler ve okulların dinamikleri farklıdır. Bu nedenle mesleki gelişimi hedefleyen etkili bir öğretmen değerlendirme modelinin bir okulun kendine ekosistemine özgü olması oldukça önemlidir (Bülbül ve diğerleri, 2013; Jones ve diğerleri, 2006; OECD, 2013a; TEDMEM, 2018; Chen ve diğerleri, 2021). Bu araştırmanın amaçları, araştırma soruları ve literatür destekli açıklamalar incelendiğinde vaka çalışmasının en uygun yöntem olacağı düşünülmüştür. Okulda yürütülen öğretmen değerlendirme uygulamalarının ancak okulun kendi bağlamında anlaşılabilceği, araştırma kapsamında geliştirilecek öğretmen değerlendirme modelinin okulun kendi kültürel yapısına özgü olmasının önemi ve katılımcıların özgün görüşlerinin bu modele yansıtılması gerektiği için vaka çalışmasının yürütülmesi uygun bulunmuştur. Vaka olarak özel bir ilkokul belirlenmiş ve araştırma 2018-2021 eğitim öğretim yılında bu

okulda gerçekleştirilmiştir. Bu okulda görev yapan sınıf öğretmenleri, yöneticiler ve uzmanlar (program geliştireme ve ölçme ve değerlendirme uzmanları) araştırmanın katılımcılarını oluşturmuştur.

2.2 Çalışmanın Bağlamı

Nitel araştırmalar bağlama bağımlı çalışmalardır. Araştırmacı, olayların, eylemlerin ve görüşlerin bağlamın benzersizliği tarafından nasıl şekillendirildiğini anlamak için bağlamsal özellikleri ve etkilerini ayrıntılı olarak tanımlamalıdır (Maxwell, 2013). Bu araştırma Ankara ilinde özel bir ilköğretim okulunda gerçekleştirilmiştir. Okulda görev yapan öğretmenlerin neredeyse tamamı en az üç yıllık öğretmenlik deneyimine sahiptir. Öğretmenler işe alınırken yazılı sınav ve mülakata tabi tutulmakta ayrıca işe alım sürecinde öğretmenlerin örnek bir ders süreci gözlemlenmektedir.

İlkokula öğrenci alım süreci de bu okulun öğretmen değerlendirme sürecini etkileyen önemli faktörlerden biridir. Araştırmanın katılımcıları, ilkokulda sınavla okumaya hak kazanan öğrencilerin anaokulundaki öğrencilere göre akademik olarak daha başarılı olduklarını düşünmektedirler. Bazı öğretmenler bu durumun öğretmenin sınıf ortalamasını ve dolayısıyla yöneticilerin öğretmenin başarısına bakışını etkilediğini bildirmiştir.

Okulda bir ilkokul müdürü ve on iki müdür yardımcısı görev yapmaktadır. Okul müdürünün 25 yıldan fazla sınıf öğretmenliği tecrübesi vardır ve müdür yardımcılarının çoğu sınıf öğretmenliği tecrübesine sahiptir. Ayrıca öğretmenler sınıfının bulunduğu her koridorda bir müdür yardımcısı bulunmaktadır ve öğretmenler bu müdür yardımcısıyla iş birliği içerisinde çalışmaktadır. Öğretmenler yöneticilerin sınıf öğretmenliği tecrübesinin olması ve iş birliği içinde çalışabildikleri gerekçisiyle bu kişiler tarafından değerlendirilmenin etkili olacağını düşünmektedirler.

Okulda beş program geliştirme uzmanı ve üç ölçme uzmanı görev yapmaktadır. Uzmanlar, ders planları, öğretim süreçleri ve değerlendirme süreçleri tasarlamak için öğretmenlerle iş birliği içinde çalışmaktadır. Uzmanlar, ayrıca projelerin uygulanmasından ve zaman zaman hizmet içi eğitimlerin düzenlenmesinden de sorumludur.

Okulda mesleki gelişim faaliyetleri genellikle hizmet içi eğitim dönemleri kapsamında yürütülmektedir. Öğretmenlere hizmet içi eğitim kapsamında hangi eğitimleri almak istedikleri sorulmakta ve öğretmenlerin bireysel istekleri yerine çoğunluğun belirlediği konularda eğitimler verilmektedir. Öğretmenlerin eğitim sonrasında öğrendiklerini nasıl uygulayabilecekleri belirlenmemektedir. Bu okulda bugüne kadar sistematik bir öğretmen değerlendirme süreci uygulanmamıştır. Ancak okul müdürü yılda bir kez her öğretmenin performansını sistematik olmayan bir şekilde gözlemlemeye çalışmaktadır.

2.3 Örneklem

Örnek olay çalışmalarında, erişilebilir olan ve araştırma sorusunu aydınlatan en önemli durumun seçilmesi çok önemlidir (Creswell ve Poth, 2018; Yin, 2018). Amaçlı örneklemede, çalışılan olguyu deneyimlemiş bireylerin seçilmesi esastır (Creswell, 2015). Bu çalışmada amaçlı örnekleme yöntemi kullanılmış olup, araştırmanın amaçları, araştırma soruları ve erişilebilirlik dikkate alınarak örnek olay seçilmiştir. Devlet okulu yerine özel okulun tercih edilmesinin nedeni, devlet okullarının öğretmen değerlendirme sürecinde okulların bireysel tercihlerinin uygulanmasından ziyade Milli Eğitim Bakanlığı'nın aldığı kararlara tabi olmasıdır. Ankara'da yer alan özel kurumlar arasından bu kurumun örnek olay olarak seçilmesinde kurumda öğretmen değerlendirme modeli uygulanmaması ve kurumun bir model geliştirmeye gönüllü olarak katılmak istemesi de belirleyici olmuştur. Okulun araştırmaya gönüllü katılımının yanı sıra, okulun araştırmacı tarafından erişilebilir olması, araştırmacının araştırmanın uzun soluklu ve detaylı bir araştırma süreci yürütebilmesi açısından da belirleyicidir. Öğretmen değerlendirmesi konusunun öğretmenlerin şüphe ve tedirginlikle yaklaştıkları bir konu olduğundan katılımcıların araştırmacıya güvenmeleri araştırmayı daha verimli ve etkili bir şekilde yürütebilmesini de sağlamıştır. Araştırmacının seçilen bu okulda program geliştirme uzmanı olarak uzun süre öğretmen ve yöneticilerle iş birliği içinde çalışması, katılımcıların soruları güvenle ve ayrıntılı olarak cevaplamalarını sağlamıştır.

Katılımcılar olarak sınıf öğretmenleri, yöneticiler ve uzmanlar (program geliştirme uzmanları, ölçme ve değerlendirme uzmanları) seçilmiştir. İlk veri toplama, mevcut

öğretmen değerlendirme sürecini, sorunlarını ve güçlü yönlerini anlamak ve bu araştırmanın ürünü olarak geliştirilecek olan değerlendirme modeli hakkında öneriler almak için yirmi sınıf öğretmeni, dokuz müdür ve yedi uzman ile gerçekleştirilmiştir. 20 öğretmen seçilirken, her sınıf düzeyinden farklı deneyim yıllarına sahip öğretmenlerin araştırmaya katılmaları sağlanmıştır. Öte yandan yöneticiler ve uzmanlardan gönüllü olarak katılmak isteyenler belirlenmiştir.

İkinci veri toplama, ilk görüşme sonuçlarına göre hazırlanan taslak model hakkında katılımcıların görüşlerini anlamak için odak grup görüşmeleri yoluyla gerçekleştirilmiştir. Odak grup görüşmelerinde katılımcıları seçerken, durumdan etkilenen, görüşlerini ayrıntılı bir şekilde ifade ederek katkıda bulunabilecek deneyime sahip ve konuyu çok yönlü anlamak için çeşitlilik sağlayabilecek kişilerin seçilmesi esastır (Bader ve Rossi, 2002; Krueger ve Casey, 2015). Ayrıca, bir odak grup katılımcı sayısı ondan fazla olmamalıdır ve altı ila sekiz katılımcıdan oluşan bir grup, içgörülerin paylaşılmasına ve tüm algıların sağlanmasına izin vermek için yeterli olacaktır (Hennink, 2014; Ledy & Ormrod, 2016). Bu çalışmanın odak grup görüşmelerinde öğretmenler, yöneticiler ve uzmanlar her bir odak grup toplantısında bir araya gelmiştir. Ayrıca hem ilk veri toplama sürecine katılan öğretmenler hem de görüşme yapılmayan öğretmenler bir araya gelerek taslak modeline ilişkin görüşlerini paylaşmış ve nelerin eklenebileceği ya da değişebileceği konusunda önerilerde bulunmuştur. Önerilen öğretmen değerlendirme modelinin geliştirilmesinde farklı deneyimlere sahip katılımcılar ile bu modelin kullanımından etkilenebilecek çeşitli görev ve sorumluluklara sahip katılımcıların dahil edilmesi, odak grup toplantılarında etkileşimi artırmış ve araştırmacının daha derinlemesine görüş almasını sağlamıştır. Tartışmaların etkili olabilmesi için gruptaki katılımcı sayısı yedi ile sınırlandırılmıştır. Görüşmeler dört farklı grup ile yapılmıştır.

2.4 Veri Toplama Araçları

Araştırmanın birinci bölümünde yer alan araştırma sorularını yanıtlamak amacıyla öğretmenlerden, yöneticilerden ve uzmanlardan veri toplamak amacıyla görüşme formları geliştirilmiştir. Görüşme formunun geliştirilme süreci üç aşamada gerçekleşmiştir. Bunlar “soruların yazımı”, “uzmanların görüşü alma” ve “pilot

uygulama” dır. Öğretmenler, müdürler ve uzmanlar için geliştirilen görüşme formları soruların kapsamı açısından benzerlik göstermektedir.

Bu çalışmada, katılımcıların değerlendirilecek öğretmen nitelikleri, zamanlama, süreç, sonuçların kullanımı ve taslak öğretmen değerlendirme modelinin uygulanabilirliği ile ilgili görüşlerini almak için dört farklı grupta odak grup görüşmeleri yapılmıştır. Odak grup görüşme formu geliştirme süreci üç adımda gerçekleştirilmiştir. Bunlar “soruların yazımı”, “uzmanların görüşü alma” ve “pilot uygulama” dır. Odak grup görüşmesinin giriş bölümünde araştırmacı bir açılış konuşması yapmış, taslak modeli sunmuş ve her gruptaki katılımcıları tanıtmıştır. Görüşme formun ayrıca, taslak modelin tüm boyutlarına yönelik hazırlanmış açık uçlu sorular içermektedir.

2.5 Araştırma Süreci

Bu araştırma üç bölümden oluşmaktadır. Araştırmanın ilk bölümü, sınıf öğretmenleri, müdürler ve uzmanların özel okulda mevcut öğretmen değerlendirme ve mesleki gelişim uygulamalarına ilişkin algılarının derinlemesine bir analizini yapmayı ve katılımcıların önerilerini derinlemesine incelemeyi amaçlamıştır. Araştırmanın ilk bölümünde birebir görüşmeler yoluyla veriler toplanmış ve veriler içerik analizi ile çözümlenmiştir. Araştırmanın ikinci bölümünde araştırmacı, çalışmanın birinci bölümünden elde ettiği görüşler doğrultusunda taslak bir model geliştirmiştir. Bu taslak model katılımcılara tanıtılmış ve odak grup görüşmeleri ile katılımcıların taslak model hakkındaki görüşleri alınmıştır. Çalışmanın üçüncü bölümünde, odak grup görüşmeleri sonrasında alınan kararlar doğrultusunda modelin kılavuzu ve modelde kullanılacak veri toplama araçları geliştirilmiştir. Kılavuz ve tüm veri toplama araçları, tüm sınıf öğretmenleri, bölüm başkanları, matematik ve fen bilgisi öğretmenleri, yöneticiler ve okulda görev yapan uzmanlarla e-posta yoluyla paylaşılmıştır ve katılımcıların görüşleri alınmıştır.

Araştırmanın birinci bölümünde görüşmeler yapılmadan önce araştırmacı, bu araştırmayı öğretmenler, uzmanlar ve yöneticiler birlikte yürütmek için okul müdürü ve kurumun genel müdüründen gerekli izinleri almıştır. Her görüşme oturumunun başında araştırmacı, katılımcıları müdür ve genel müdürün bu araştırmadan haberdar

olduđu konusunda bilgilendirmiştir. Bu açıklama, gizli veya yasa dışı bir işlem yürüttükleri izlenimi vermek yerine kendilerini rahat hissetmelerini ve soruları içtenlikle yanıtlamalarını sağlamıştır. Ayrıca araştırmacı, görüşmeleri sohbet şeklinde gerçekleştirmiş ve bu görüşmelerin yalnızca araştırma amaçlı yapıldığını, gizlilik ilkesine tamamen uyulacağını açıklayarak katılımcıların içtenlikle yanıt verebileceklerini belirtmiştir. Katılımcıların olumsuz deneyimleri açıklamak istemediğini fark edildiğinde araştırmacı kendilerini rahat hissetmezlerse açıklama yapmak zorunda olmadıklarını açıklamıştır. Katılımcılar, araştırmacıya ve araştırmanın kendisine güvendiklerinden bahsetmek istemedikleri süreçler hakkında da görüşlerini açıklamışlardır ve bu sayede derinlemesine ve çok yönlü veriler elde edilmesi mümkün olmuştur. Görüşmeler, katılımcıların izni alındıktan sonra ses kaydına alınmıştır. Araştırmacı, katılımcılara görüşmenin herhangi bir aşamasında kendilerini rahatsız hissettiklerinde görüşme oturumunu derhal sonlandıracaklarını ve kalan verilerin silineceğini de açıklamıştır. Tüm katılımcılara soruları yanıtlamaları için yeterli zaman verilmiştir. Her görüşmeden hemen sonra kısa bir izleme görüşmesi yapılmıştır. Bu görüşmelerde araştırmacı tarafından alınan notlar ve her bir soruya verilen yanıtlar katılımcıya gerek duyduğu yerleri açıklığa kavuşturması veya eklemeler yapmasına fırsat vermesi için yönlendirilmiştir. Bire bir görüşmeler öğretmenler için yaklaşık kırk dakika, müdürler için otuz dakika ve uzmanlar için kırk beş dakika sürmüştür.

Odak grup görüşmesinde araştırmacı moderatör olarak soruları sormuş, herhangi bir yargıda bulunmamış, katılımcıları etkili bir şekilde dinlemiş ve herkesin eşit konuşma şansına sahip olmasını sağlamıştır. Odak grup görüşmelerinin bir ders saatinden uzun sürebileceği gerekçesiyle görüşmeler, öğretmenlerin ders saatleri dışında kalan toplantı saatlerinde gerçekleştirilmiştir. Odak grup görüşmeleri altmış beş ile seksen beş dakika arasında sürmüştür. Tüm katılımcılardan izin alındıktan sonra ses kaydı alınmıştır. Odak grup tartışmasının hemen ardından kısa bir izleme görüşmesi gerçekleştirilmiştir. İzleme görüşmesinde, araştırmacı tarafından alınan notlar ve her bir soruya verilen yanıtlar, tüm katılımcılara bazı noktaları açıklığa kavuşturmak ve eğer ilave açıklamalar eklemek isterlerse ekleme yapma şansı vermek için yönlendirilmiştir.

Üçüncü veri toplama süreci, katılımcıların modelin son hali hakkında görüşlerini almak ve değerlendirme modelinin veri toplaması için hazırlanan araçları doğrulamak için görüşlerini almak üzere gerçekleştirilmiştir. Model açıklamalarını ve veri toplama araçlarını içeren kılavuz, tüm katılımcılara görüşlerini almak üzere e-posta yoluyla gönderilmiştir. Katılımcıların rehber ve öğretmen değerlendirme araçlarına ilişkin görüşleri çevrimiçi bir geri bildirim formu aracılığıyla toplanmıştır. Form, araştırmanın amacı ve sürecinin ayrıntılı bir açıklamasıyla birlikte tüm sınıf öğretmenlerine, matematik ve fen öğretmenlerine, müdürlere ve okulda görev yapan uzmanlara e-posta yoluyla gönderilmiştir.

2.6 Veri Analizi

Bu çalışmada veriler içerik analizi yardımıyla incelenmiştir. İçerik analizinin ilk adımı, görüşme kayıtları dinlenerek görüşme dökümlerinin oluşturulmasıdır. Araştırmacı, ses kayıtlarını yazıya dökmüş ve yazıya dökülen metin araştırmacı tarafından satır satır okunmuştur. Bu süreç, araştırmacının yazıya aktarılan verilere aşina olmasına ve kodlar arasındaki ilişki kurarak tema ve kategoriler üretmesine yardımcı olmuştur. Bu çalışmada okula özgü bir öğretmen değerlendirme modeli geliştirileceği gerekçesi ile birçok farklı görüşe ve olasılığa ihtiyaç duyulmuştur. Bu amaçla hem bire bir hem de odak grup görüşmeleri için açık kodlama süreci gerçekleştirilmiştir. Araştırmacı, bireysel görüşmelerin ses kayıtlarını yazıya geçirdikten sonra, verileri birkaç kez baştan sona okumuş ve manuel olarak kodlamıştır. Yazıya dökülen metin ilk kodlama döngüsü için analiz edilmiş ve anlamlı olan kelime veya cümleler belirlenmiştir. Araştırmacı, yazıya aktarılan verilerin sağ kenarına kodları not etmiştir. İkinci kodlama döngüsünde kodlar arasındaki ilişki incelenmiş ve benzer boyutu temsil eden kodlara aynı kategoriler altında yer verilmiştir.

2.7 Araştırmanın Geçerliği

Nitel araştırmalarda geçerlik toplanan verilerin gerçeği yansıtmasıyla ve bulguların gerçeğe ne kadar uyumlu olduğuyla ilgilidir. Çeşitli kişilerden birden fazla yöntem kullanarak bilgi toplanması yanlışlık riskini azaltır (Maxwell, 2013). Araştırmacının

araştırdığı olguyu olduğu biçimiyle ve olabildiğince yansız gözlemlemesi geçerliği doğrudan etkiler (Kirk ve Miller, 1986). Nitel araştırmalarda iç geçerliği sağlanmasında kullanılan bir yöntem veri kaynağında ve/veya veri toplam araçlarında çeşitliliğin sağlanmasıdır. Bu araştırmada toplanan veriler, bu okulda çeşitli alanlarda (öğretmenler, yöneticiler ve uzmanlar) çalışan farklı bakış açılarına sahip katılımcıların görüşlerini yansıtmaktadır.

Uzun süreli etkileşim sağlayarak geçerli veriye ulaşmakta iç geçerlik için önemlidir. Uzun süreli etkileşim için de kültürü öğrenmek, katılımcılarla güven ve sağlam ilişkiler kurmak ve katılımcıların açıklamaları için açık ve kapsamlı olmak için yeterli zaman harcamak esastır (Patton, 2015). Araştırmacı uzun yıllardır araştırmanın yürütüldüğü kurumda çalışmıştır ve tüm katılımcılarla yeterince zaman geçirmiştir. Araştırmacının okul kültürünü bilmesi ve katılımcıların çoğuyla daha önce çalışmış olması, katılımcıların araştırmacıya güvenmelerini ve yaptıkları açıklamaları gerçeği yansıtmalarını sağlamıştır.

Uzman incelemesi araştırmacının vardığı sonuçların verilere dayandığından emin olmak için kullanılır (Creswell & Poth, 2018). Diğer prosedürler arasında, uzman incelemesi, geçerlik için en uygun yöntem olarak tanımlanabilir (Patton, 2015). Uzman incelemesi, araştırma sürecinin nesnel bir yaklaşımla yürütülüp yürütülmediğinden emin olmak için araştırmanın veri toplama, veri analizi ve raporlama süreçlerinde uzmanlara danışılması ve araştırmacının araştırmanın birçok aşamasında varsayımlar, önyargılar ve yorumlar konusunda kendini sorgulamasını sağlar (Ravitch & Mittenfelner, 2016). Bu çalışmada araştırmacı, uzman görüşü almak için her bir araştırma aşamasında, araştırmacının tez danışmanına ve tez izleme komite üyesi olan ve nitel araştırma çalışmalarında deneyimli iki araştırmacıya danışmıştır.

Katılımcı teyidi ile çalışmanın bulguları katılımcılarla paylaşılır ve katılımcıların eleştiri veya geri bildirim sağlayarak yorumlar yapmasına fırsat verilir. Bu sayede veri analizi aşamalarında katılımcıların bakış açılarına yer verilmesine olanak sağlanır (Creswell, 2015; Tracy, 2020). Katılımcı teyidi, yanıtların yanlış yorumlama olasılığını ortadan kaldırmanın en önemli yoludur ve katılımcıların bakış açıları yanıtlayıcı doğrulaması için kullanılır (Maxwell, 2013). Her görüşmeden sonra

araştırmacı, doğrulama almak, yanlış anlamaları düzeltmek veya ek bilgi sunmak için görüşülen katılımcılara yanıtlarını özetlemiştir. Ayrıca görüşmelerden elde edilen verilere dayalı oluşturulan taslak öğretmen değerlendirme modeli ve modelin son hali, katılımcılarla paylaşarak görüşleri teyit edilmiştir.

Aktarılabilirlik, okuyucunun çalışılan durum ile bulguların aktarılabilceği durum arasındaki benzerlik derecesini belirleyebilmesi için okuyucuya incelenen vaka hakkında yeterli bilgi sağlamak ile ilgilidir (Patton, 2015). Bu çalışmada, bu araştırmadan elde edilen bulguların benzer bağlamda başka bir vakaya uygulanabilmesi için “Çalışmanın Bağlamı” başlığı altında ayrıntılı olarak vakaya ilişkin bir betimleme yapılmıştır. Araştırmacı, çalışmanın bulgularını rapor ederken, okuyucuların bağlamı ve katılımcıların duygu ve algılarını anlamaları için alıntılar da kullanmıştır.

2.8 Araştırmanın Güvenirliği

Pozitivist tekniklerin kullanıldığı nicel araştırmalarda, çalışma aynı bağlamda, aynı yöntemlerle ve aynı katılımcılarla tekrarlanırsa güvenilirlikten bahsedilebilir. Nitel bir çalışmada bir araştırmacı, çalışma içindeki süreç ayrıntılı olarak rapor ederse çalışmayı tekrar edebilir, bu da okuyucunun nitel araştırma basamaklarının yeterince takip edilip edilmediğini değerlendirmesine olanak tanır (Shenton, 2004). Güvenilirliği sağlamak için, doktora komitesi üyeleri gibi dış denetleyiciler, veri toplama ve veri analiz sürecinin kalitesi hakkında yargıda bulunabilir (Patton, 2015). Bu araştırmada hem veri toplama süreci hem de veri analiz süreci araştırmacının tez danışmanı, tez izleme komitesi üyeleri ve nitel araştırma alanında çalışan iki araştırmacı tarafından detaylı olarak incelenmiştir. Araştırmacı, elde edilen geri bildirimler ve katkılarla araştırma sürecini şekillendirmiştir.

Objektiflik olarak da kabul edilen teyit edilebilirlik, bir sorgulamanın verilerinin ve yorumlarının araştırmacının hayal gücünün ürünü olmadığını kanıtlanması ile ilgilidir (Patton, 2015). Doğrulanabilirlik, araştırmacı yanlılığının minimum düzeyde tutulması ve sonuçların, araştırmacının araştırmasından ziyade araştırmanın konusuna ve koşullarına bağlı kılınmasıyla ilgilidir (Miles ve Huberman, 1994).

Doğrulanabilirliği sağlamanın bir yolu, çalışmanın her aşamasının güvenilir bir dış gözlemci tarafından ayrıntılı olarak incelenmesidir (Lincoln ve Guba, 1985). Bu araştırmanın genel yöntem ve prosedürlerinin açık ve ayrıntılı olarak açıklanıp açıklanmadığı, verilerin nasıl toplandığı, işlendiği ve dönüştürüldüğü, bulguların araştırma ile net bir şekilde ilişkili olup olmadığı gibi önemli konularda tez danışmanından sürekli geri bildirim alınmıştır.

Araştırmacı, kişisel varsayımlar, değerler ve önyargılar, duygusal durumlar ve bunların çalışma sırasında nasıl ortaya çıkmış olabileceği konusunda mümkün olduğunca açık ve öz farkındalığa sahip olmalıdır (Miles ve diğerleri, 2014). Bu amaçla araştırma sürecinde yansıtıcı açıklamalar şeklinde notlar eklenebilir (Ortlipp, 2015). Doğrulanabilirliği sağlamak için, araştırmacı, görüşmelerin yapıldığı her günün ardından notlar tutmuştur ve her veri toplama oturumunun izlenimlerini, potansiyel önyargıları ve araştırma sürecini ve sonuçlarını etkileyebilecek yatkınlıkları kaydetmiştir. Araştırmacı bu notlardan özellikle veri analizi sürecinde ve sonuçların raporlanması bölümünde yararlanmıştır.

BULGULAR

3.1 Öğretmen Değerlendirmesine İlişkin Mevcut Uygulamalar ve Sorunlar

Katılımcılar araştırmanın yürütüldüğü bu kurumda değerlendirmenin amacının net olmadığını, değerlendirme için belirli bir planın veya zaman çizelgesinin yer almadığını, geçerliği sağlanmış standart formların kullanılmadığını ve değerlendirmelerin resmi olmayan okul müdürü gözlemlerinin yapıldığını belirtmişlerdir. Katılımcılar 2018 yılında ilk kez müdürün habersiz bir şekilde informal gözlemler yaptığını belirtmişlerdir. Bulgular bu süreçte okul müdürünün sınıfın arkasında oturarak notlar tuttuğu ancak bir form kullanmadığı ve yapılan gözlemlerde hiçbir ölçüt ya da standardın kullanılmadığını bu nedenle de katılımcıların gözlemlerin ne amaçla yapıldığını ve hangi davranışlarının değerlendirildiğini bilmediklerini ortaya koymuştur.

Katılımcılara, değerlendirme amacıyla veri toplama sürecinin okulda nasıl yürütüldüğü sorulduğunda, tüm öğretmenler, müdürler ve uzmanlar, öğretmen

değerlendirmesinde veri kaynağı olarak sadece akademik başarıya dayalı okul genelinde yapılan sınavlardan elde edilen puanların kullanıldığını belirtmişlerdir. Katılımcılar, bu puanların öğretmenleri değerlendirmek için nasıl kullanıldığı sorduğunda, okul genelinde yapılan her sınavdan sonra tüm öğretmenlere kendi sınıftaki öğrencilerin ortalama puanını içeren bir raporun sunulduğunu belirtmişlerdir. Bu raporlara göre, öğretmenlerin sınıflarının başarısı, öğretmenin başarısına eşdeğer kabul edilmektedir. Bu bulgudan yola çıkılarak katılımcılara değerlendirme sonuçlarının ne amaçla kullanıldığı sorulduğunda yazılı olarak açıklanmış bir değerlendirme amacı olmamasına rağmen değerlendirmelerin genel olarak öğrenci başarısını arttırmak amacıyla yapıldığı açıklanmıştır.

Literatür incelendiğinde değerlendirmelerin yasal, etik olarak yapılmalı ve değerlendirmeye dahil olan katılımcıların refahı göz önünde bulundurulmalıdır (Howard ve Gullickson, 2009). Değerlendirme sisteminin amacını, sistemin nasıl kullanılacağını, veri toplama ve raporlama sürecini açıklayan politika beyanları ve/veya kılavuzlar, değerlendirmelerin tutarlı, adil ve eşitlikçi olmasını sağlar. Bu okulda değerlendirmelerin açıklandığı prosedür ya da politikaların olmayışı katılımcılarda değerlendirmelerin eşit, adil ve tutarlı olması konusunda şüphe uyandırdığını bulunmuştur. Yöneticiler ve uzmanlar da bu durumun öğretmenlerin değerlendirme sürecinin uygunluğu konusunda endişe duymalarına neden olduğunu ve bu nedenle kendilerini her zaman tedirgin hissettiklerini belirtmişlerdir. Bir yönetici (Y2), bu durumu şöyle açıklamaktadır.

Koridorlarda bazen gözlem yapmaya çalışıyorum. Koridordan geçerek öğretmenin sınıfına kapıdan bakıyorum. O anda ne görebildiysem bir fikir edinmeye çalışıyorum. Ama aslında, yönetici olarak ne sıklıkla gözlemlemem gerektiğini bilmiyorum. Hangi niteliklere bakmam gerektiğini bilmiyorum. Sadece sınıfı kapısından yapılan iki dakikalık gözlemle doğru ya da adil değerlendirme yapmak mümkün değil ki. Bir kere etik de değil.

Öğretmenlerin çoğunluğu gözlemlerden önce ve sonra bilgilendirilmediklerini; ayrıca, değerlendirme sonuçları hakkında yöneticilerden hiçbir zaman bireysel geri bildirim almadıklarını belirtmişlerdir. Bu nedenle, gözlemlenen veya değerlendirilen zayıf veya güçlü yönlerden haberdar olmadıklarını belirtmişlerdir. Öğretmenler, sınıf

gözlemleri veya herhangi bir değerlendirmeden elde ettikleri ve farkında olmadıkları bilgilerin başka nedenlerle kullanıldığını düşündükleri için kendilerini rahatsız hissetmektedirler. Öğretmenlerin çoğu geribildirimlerinin dengeli olmadığını ve genel olarak zayıf yönleriyle ilgili bireysel geribildirim aldıklarını belirtmişlerdir. Öğretmenler özellikle okul genelinde yapılan sınavlarda sınıf ortalamaları düştüğünde, okul genelinde sorunlar yaşandığında ve veliler şikâyet ettiklerinde bire bir geri bildirim aldıklarını belirtmişlerdir. Deneyimli öğretmenlerden biri (Ö11) bu durumu şu şekilde açıklamıştır:

Deneyimli bir öğretmenim ve müdürüm beni her zaman takdir eder, ancak müdür bir başka öğretmeni aradığında ve onunla görüşmeyi istediğinde öğretmenlerin çoğu endişeleniyor çünkü yöneticiler bireysel olarak konuşmak istediklerinde sadece olumsuz gördüğü şeyleri söylüyor. Takdir edilen bir öğretmen olsam bile, bunu doğru bulmuyorum. Keşke öğretmenlere iyi yaptıkları şeyler hakkında da sık sık bilgi verilse.

Değerlendirmelerin değerlendirilen kişiye faydalı olabilmesi için bilgilendirici olması ve zamanında verilmesi gereklidir (Howard ve Gullickson, 2009). Sonuçların kullanılabilirliği için, değerlendirme sistemleri iyi tanımlanmış ve açık ölçütler kullanarak performans beklentilerini özetlemelidir. Ayrıca değerlendirmelerde, performansı net ve savunulabilir bir gerekçeye dayalı olarak yorumlamak veya yargılamak için iyi tanımlanmış ölçütler kullanılmalıdır. Aksi takdirde bireysel yorumlar yapılabilir ve bu yorumlar sonuçların kullanılabilirliğini tehlikeye atar. Bu örnek olayda katılımcılar açık ölçütlerin olmamasını bir sorun olarak belirtmiştir. Katılımcılar değerlendirmede savunulabilirliğin esas olduğunu ve bunun açık ölçütlerle sağlanabileceğini açıklamışlardır.

Değerlendirmelerin etkili ve kullanışlı olması için, sonuçların detaylı bir şekilde yer aldığı işlevsel raporlara ihtiyaç vardır ve bu raporlar zamanında hazırlanmalıdır ki öğretmenler belirtilen eylemleri gerçekleştirebilsin (Howard ve Gullickson, 2009). Bu örnek olayda katılımcılar zamanında verilmiş yazılı bir rapor olmadığını belirtmişlerdir. Katılımcılar bunu bir sorun olarak tanımlamışlardır, çünkü onlara göre raporlama, güçlü yönlerin daha da geliştirilmesi ve zayıf yönlerin iyileştirilmesi için esastır. Ayrıca mesleki gelişim için geri bildirim içeren raporlama, öğretmenlerin bir

plan geliřtirmesine olanak tanır ve raporlama süreci geliřimin kayıt altında tutulmasına yardımcı olur. Bir öęretmen (Ö8) bu durumu řöyle açıklamaktadır.

řimdi, birisi bana ne tür eęitimlere ihtiya duyduęumu sorsa, sadece kendi deęerlendirmem doęrultusunda açıklayabilirim. Zayıf yönlerimle ilgili geri bildirimler yazılı bir raporla düzenli aralıklarla sunulsaydı bu soruya daha doęru yanıtlarım.

Doęruluk terimi, deęerlendirmeden elde edilen verilerin karar vermek ve yargıda bulunmak için geçerli olup olmadıęını belirlemek ile ilgilidir (Howard ve Gullickson, 2009). Deęerlendirme yapılırken birden fazla veri kaynaęı kullanılarak sonuçların yanlış yorumlanmasının önüne geçilebilir (Howard ve Gullickson, 2009). Katılımcılar öęretmenler arasında oluřturulan rekabet ortamı ve gözlem sürecinin oluřturabileceęi aksaklıklar nedeniyle öęretmenin etkililięinin yanlış yorumlandıęı belirtilmiřtir. Katılımcılar sınıf ortalama puanlarının dięer öęretmenler ile kıyaslanmasının ve bu kıyaslama sonucunun daha başarılı ya da daha az başarılı olarak yorumlanmasının öęretmenin gerçek performansını yansıtmadıęını açıklamıřlardır. Bir öęretmen (Ö1) rekabet ortamının oluřturduęu problemi řöyle ifade etmiřtir.

Herkes birbiriyle kıyaslanıyor. Başarılı olmak öęretmenler arasında ego savařı gibi. Sınıfınızda önemli bir başarı elde ettiyseniz, bu sınıfınızın başarısıdır. Başkalarıyla karřılařtırılmamalı. Benim hakkımda başkalarıyla kıyaslama yapılarak karar verilmesini doęru bulmuyorum.

Katılımcılar özellikle de öęretmenler arařtırmanın yürütüldüęü bu okulda tecrübeli öęretmenlerin lehine ve bireysel iliřkilere göre deęerlendirme yapılmasının deęerlendirme sonuçlarında yanlılıęa sebep olduęunu ve dolayısıyla da sonuçların doęruluęunu negatif etkiledięini belirtmiřlerdir. Bu kurum için tecrübeli öęretmenlerin hem sınıf içi hem de sınıf dıřında gerçekleřtirdikleri faaliyetlerde olumlu geribildirim aldıkları, eleřtirilmedikleri ve başarılı olarak deęerlendirildikleri söylenebilir. Bir öęretmen (Ö19) önyargılı deęerlendirmelerin oluřturduęu problemi řöyle ifade etmiřtir.

oęu zaman, katıldıęım eęitimlerden öęrendięim öęretimsel yöntemleri ya da materyalleri kullanmak istedięimde, zümremdeki tecrübeli öęretmenler "bu okulda iře yaramaz, deneyimsizsiniz tabi bilmiyorsunuz" diyorlar. Bu sefer sınıfımda ısrar edip uyguladıęımda,

sınıfta gürültü oluyor diye yöneticilerden gelen uyarılarla karşı karşıya kalıyorum. Öte yandan, deneyimli öğretmenler yirmi yıl önce bildiklerini uygulamaya devam ediyorlar ve daha hiç uyarı aldıklarını görmedim.

Bu okulda öğretmenlerin sınıf içinde gerçekleştirdikleri öğretimi değerlendirmek amacıyla ders planlarının ve okul geneli sınavlardan elde edilen ortalamalarının kullanımının geçerli veri sağlamada önemli bir problem olduğu ve doğruluğu olumsuz etkilediği tespit edilmiştir. Katılımcılar sınıf içi gözlemler yapılmadan öğretimlerinin etkililiği hakkında yorum yapılmasını doğru bulmamaktadır. Öğretmenler ders planlarının öğretme/öğrenme sürecinin karmaşıklığını yansıtmadığını ve hazırladıkları ders planlarının niteliğine ilişkin herhangi bir geri bildirim almadıklarını da belirtmişlerdir. Ayrıca öğretmenler ve uzmanlar, sadece bilişsel boyutun ölçüldüğü okul geneli sınavlardan elde edilen ortalama puanlar ile yapılan değerlendirmelerin, duyuşsal boyutta öğretmenin ne kadar etkili olduğunun gözden kaçırılmasına sebep olduğunu belirtmişlerdir. Bir öğretmen (Ö15) bu problemi şöyle ifade etmiştir.

Sadece bilişsel alana göre değerlendirilmek, sınıflarımızdaki bazı çocukları kaybettiğimiz anlamına gelir. Özel ihtiyaçları olan çocuklar ne olacak? Sınav sonuçlarına göre yönetimin her öğretmen hakkında genel bir görüşü var ve aslında bir öğretmen imajı oluşmuş durumda. Daha fazla akademik başarı getiren öğretmen iyi öğretmendir. Neden sadece akademik başarıları takip ediyoruz? Bakalım öğrencilerin duygusal yönlerini destekliyor muyuz?

3.2 İhtiyaç Duyulan Öğretmen Değerlendirme Modeli

Katılımcılar modelin amacının öğretmen yeterliklerinin (iletişin ve iş birliği, öğretim, okula karşı sorumluluklar, derse hazırlık, izleme ve değerlendirme, mesleki gelişim) değerlendirilmesi ve geliştirilmesi olması gerektiğini belirtmişlerdir. Öğretmenleri yöneticilerinin, sınıf öğretmenlerinin, farklı branştan öğretmenlerin ve öğretmenin kendisinin değerlendirebileceği ve bu değerlendiricilerin adil, olumlu iletişim becerilerine sahip, sınıf öğretmenliği yapmış, empati kurabilen, öğretmeni değerlendirme konusunda yetkin, okul hakkında bilgi sahibi ve etkili iletişim kurabilen kişiler olmasının öneminden bahsetmişlerdir. Katılımcılar ayrıca değerlendiricilerin modelinin uygulanmasından önce veri toplama araçlarını nasıl kullanacakları ve

değerlendirme sürecinde nelere dikkat etmeleri gerektiği ile ilgili eğitim almaları gerektiğini belirtmişlerdir.

Katılımcılar gözlemlerle ilgili sorunların okulda sistematik gözlemlerin yapılmaması, öğretmenlerin gözlem sonuçları hakkında bilgi sahibi olmaması ve değerlendirme kriterlerinin olmaması olduğunu belirtmişlerdir. Bu nedenle katılımcılar değerlendirmede veri toplama aracı olarak belli ölçütler doğrultusunda hazırlanmış geçerli gözlem formlarıyla yapılan değerlendirmelere ihtiyaç olduğunu belirtmişlerdir. Ayrıca gözlemin öncesinde ve sonrasında yapılan görüşmelere de ihtiyaç duyulduğu açıklanmıştır. Bir öğretmen (T4) gözlemin önemini şöyle ifade etmiştir.

Gözlemin çok önemli olduğunu düşünüyorum ve sınıfımı birçok kez gözlemlenmesinden mutluluk duyarım. Sınavlarda başarısı düşük olan öğrencilerin raporları sürekli olarak bizlere gönderiliyor ve bu çocuklar için sanki hiçbir şey yapmadığımız düşünülüyor. Sınıfıma gelip onlara ne kadar emek harcadığımı görmelerini çok isterim. Bunu yaparken, hangi kriterlere bakacaklarını ve ne sıklıkla gözlem yapacaklarını bizimle paylaşmaları iyi olacaktır.

Katılımcılar değerlendirme sonuçlarının gizlilik ilkesine uygun olarak sadece öğretmenin kendisiyle paylaşılması gerektiğini belirtmişlerdir. Katılımcılar ayrıca değerlendirme sonuçları hem güçlü hem de zayıf yönleri içermesi gerektiğini, geri bildirimleri ve çoklu ölçümlere dayalı öğretmen gelişim planını içeren bir raporun işlevsel olacağını ve öğretmenlere gelişimleri için yeterli zaman tanımak için her eğitim döneminin sonunda raporun verilmesi gerektiğini açıklamışlardır.

Değerlendirme sonuçlarının önyargısız olması için verilerin objektif olarak elde edilmemesi yani kişisel ilişkilere dayalı kararlar alınmaması gerektiği de katılımcılar tarafından vurgulanmıştır. Bu kurumda çalışan öğretmenler, deneyimli öğretmenlerin davranışlarına, bu kişilerin müdürlerle olan iyi ilişkilerine bağlı olarak müsamaha gösterildiğini belirtmişlerdir. Öte yandan yeni başlayan bir öğretmenin davranışı sürekli incelenmiş ve eksik bulunmaya çalışılmıştır. Katılımcılar mesleki deneyimleri ne olursa olsun tüm öğretmenlerin aynı süreci takip ederek değerlendirilmesinin önemli olduğunu belirtmişlerdir. Bir öğretmen (T15) bu durumu şöyle açıklamıştır.

Bu okula ilk başladığımda kıyafetlerim bile eleştirildi, ancak başka bir öğretmen kıyafeti konusunda eleştiri almadı çünkü bu okulda çok

fazla deneyime sahipti. Her konuda tecrübeli öğretmenlerin lehine bir değerlendirme söz konusu. Bu nedenle değerlendirme, öğretim deneyiminden bağımsız olarak herkes için aynı olmalıdır, hatta kriterler aynı olmalıdır. Bu durumda, gözlemcinin yapacak çok işi olduğunu ve farklı deneyime sahip öğretmenlere eşit davranması gerektiğini düşünüyorum.

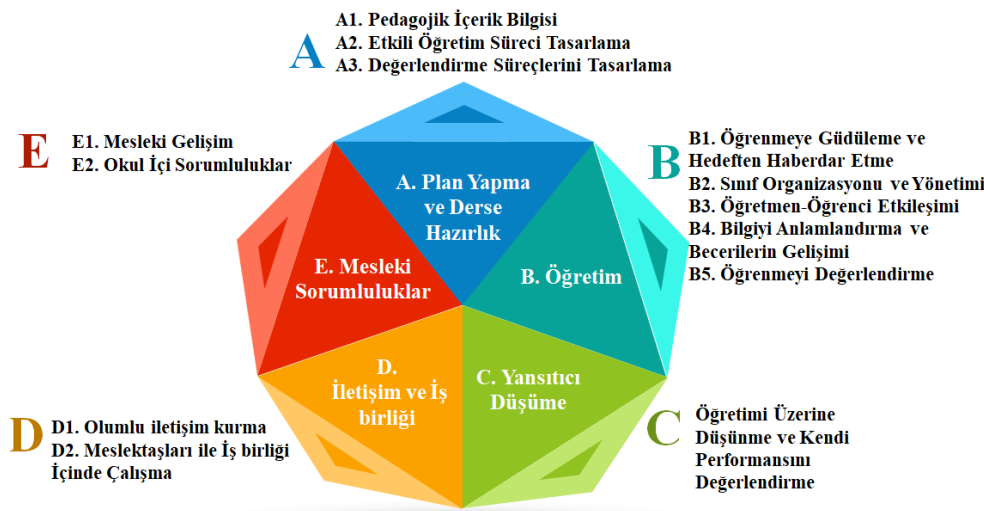
Okulda hizmet içi eğitim seminerlerinin öğretmenlerin çoğunluğunun seçtiği konular veya popüler konular üzerinden yürütüldüğü ortaya çıkmıştır. Bu eğitim genellikle tüm öğretmenlere toplu olarak verilmekte ve öğretmen nadiren aktif olarak katılmaktadır. Geniş kitlelere verilen genel eğitim, bireysel ihtiyaçları karşılamadığı, öğrenilenlerin kalıcı olmadığı ve katılımcıların katılım sağlayamadığı için etkisiz kalma sorunu olarak yorumlanmaktadır. Katılımcılar ağırlıklı olarak mesleki gelişim etkinlikleri planlanırken öğretmen değerlendirme sonuçlarının kullanılması gerektiğini öne sürmüşlerdir.

3.3 Geliştirilen Öğretmen Değerlendirme Modeli

Araştırmanın ikinci bölümünde araştırmacı, çalışmanın birinci bölümünden elde ettiği görüşler doğrultusunda bir taslak modeli geliştirmiştir. Katılımcılara taslak model tanıtılmış ve odak grup görüşmeleri ile katılımcıların görüşleri alınmıştır. Odak grup görüşmelerinin sonucunda modelde yer alması planlanan yeterlik alanları ve alt boyutları değiştirilmiş ve düzenlenmiştir (Şekil 1). Katılımcılar her bir yeterlik alanının eşit öneme sahip olduğunu ve önceliklendirilmesine gerek olmadığını ifade ederken yeni ve tecrübeli öğretmenler için bir farklılık yapılmaması gerektiği de vurgulanmıştır. Değerlendiricilere verilecek eğitimlerin uygulamalı olması ve düzenli olarak tekrarlanması da bu görüşmelerde kararlaştırılmıştır.

Daha önce yapılan görüşmeler incelendiğinde öğretmenlerin büyük çoğunluğunun ders planlarının değerlendirilmesinin süreci tam olarak yansıtmayacağı için yapılmaması gerektiğini belirlenmişlerdir. Odak grup görüşmeleri ile katılımcılara böyle bir model içerisinde ders planının kullanılması hakkında ne düşündükleri sorulmuştur. Katılımcılar ders planlarının tek bir kaynak olarak görülmediği ve bu modelde belirtilen aşamalar gerçekleştirildiği sürece değerlendirilmesinde bir sakınca olmadığını belirtmişlerdir.

Katılımcılardan bu modelin uygulanabilirliği hakkında yorum yapmaları istenmiştir. Ayrıca bu modelle değerlendirilmek mi yoksa değerlendirme yapmak mı isteyip istemediklerini sordular. Katılımcılar bu modelin uygulanması için belirtilen sürelerin uygun olduğunu belirtmişlerdir. Odak grup görüşmelerinde katılımcıların çoğu bu modelde kendi görüşlerinin yer aldığını ve işbirlikçi bir model olduğunu ifade etmişlerdir bu durumun kendilerini değerli hissettirdiğini açıklamışlardır. Katılımcılar kendi görüşlerinin yansıtıldığı ve ihtiyaçlarının dikkate alındığı bu model kapsamında değerlendirilmeye istekli olduklarını ifade etmişlerdir



Şekil1: Modelde yer alan yeterlik ve alt yeterlik alanları

Katılımcılara taslak modelin farklı boyutları ile ilgili sorular sorulduğunda ve özellikle uygulanabilirliğini değerlendirmeleri istendiğinde bu modelin daha detaylı anlatıldığı bir rehber ihtiyacı olduğunu belirtmişlerdir. Tüm gruptaki katılımcıların ortak görüşü, bu modelin amacının, yeterlilik alanlarının, kullanılacak araçların, zaman çizelgesinin, sürecin nasıl yürütüleceğinin ve öğretmenlerin nasıl bilgilendirileceğini ayrıntılı olarak açıklandığı bir rehber kitap hazırlanabileceği ve bu modelin uygulanmasında birliğin sağlanması açısından bu kitapçığın önemli olduğu yönündedir. Bu amaçla toplantı sonunda bir rehber kitapçık hazırlanmasına karar verilmiştir. Katılımcılar ayrıca veri toplama araçlarının detaylı ölçütler doğrultusunda hazırlanarak kendilerine sunulması gerektiğini, bu araçların incelenmesinde ve

modelin kendisine katkıda bulunmaya istekli olduklarını belirtmişlerdir. Birçok katılımcı, bu araçların pilot uygulama yapılması halinde daha işlevsel olacağını da eklemiştir.

Çalışmanın üçüncü bölümünde, odak grup görüşmeleri sonrasında alınan kararlar doğrultusunda modelin kılavuzu ve modelde kullanılacak veri toplama araçları geliştirilmiştir. Kılavuzda tüm yeterlik alanları, alt boyutları ve ölçütler detaylı bir biçimde verilmişti. Modelin amacı, öğretmen değerlendirmenin önemi ve bu modelin geliştirilme aşamaları da kılavuzda yer almaktadır. Model yardımıyla değerlendirmelerin nasıl yapılacağı ve objektifliğin, tutarlılığının, gizliliğin nasıl sağlanacağı da detaylı bir biçimde açıklanmıştır. Kılavuzda yer alan ölçütler yardımıyla veri toplama araçları geliştirilmiştir. Rehber ve tüm veri toplama araçları, tüm sınıf öğretmenleri, bölüm başkanları, matematik ve fen bilgisi öğretmenleri, müdürler ve okulda görev yapan uzmanlarla e-posta yoluyla paylaşılmıştır. Alınan görüşler doğrultusunda kılavuzun bazı bölümlerine eklemeler yapılmış, sözcükler düzenlenmiş ve bazı ifadeler daha anlaşılır hale getirilmiştir. Ayrıca bu bölümde, veri toplama araçlarının geçerliliği sağlamak amacıyla ön-pilot uygulamaları da gerçekleştirilmiştir.

TARTIŞMA VE ÖNERİLER

Bu çalışma, okuldaki öğretmen değerlendirmelerinin sistematik olmaktan uzak olduğunu bu okulda çalışanların, amacı açık, açıklayıcı kriterler içeren, belirli bir değerlendirme takvimi olan ve geçerli veri toplama araçlarını kullanan sistematik bir öğretmen değerlendirmesine ihtiyaç duyduklarını göstermiştir. Bu sistematik olmayan değerlendirmeler nedeniyle katılımcılar öğretmen değerlendirmeleri hakkında yeterli bilgiye sahip değildir ve yaşadıkları belirsizlik, değerlendirme sürecine güven duymalarını engellemektedir. Öğretmenlerin herhangi bir biçimlendirici, yapıcı veya bireysel geribildirim almadıkları da ortaya çıkmıştır ve öğretmenlere genel toplantılarda sadece özetleyici dönüt verildiği bu çalışmanın önemli bulgularından biridir. Bu örnek durumda, politika ve prosedürlerin eksikliği, değerlendirme sonuçlarına erişim eksikliği ve sadece eksikliklerin aktarıldığı dengeli bir geri bildirim sürecinin olmaması değerlendirmeler doğruluğunu da olumsuz etkilemiştir.

Değerlendirme sisteminin amacını, sistemin nasıl kullanılacağını ve veri toplama ve raporlama sürecini açıklayan politika beyanları ve/veya kılavuzlar, değerlendirmelerin tutarlı, adil ve hakkaniyetli olmasını sağlar (Howard & Gullickson, 2009). Farklı araştırmalarda sistematik olmayan bu tür öğretmen değerlendirmelerine ve bu durumun oluşturduğu benzer problemlere dikkat çekmiştir (Collins, 1999; Fowler, 2001; Türkoğlu, 2015; Weisberg ve diğerleri, 2009). Araştırmalar ayrıca ayrıntılı geri bildirim eksikliğinin (Sinnema ve Robinson, 2007) ve etkili biçimlendirici geri bildirim eksikliğinin (Kraft ve Gilmour 2017; Lavigne, 2014) öğretmenlerin karşılaştığı sorunlardan bazıları olduğunu göstermiştir.

Araştırmanın yapıldığı okulda değerlendirme sonuçlarının ve geri bildirimlerin paylaşılmasına yönelik açık kriterlerin ve işlevsel bir raporlama sürecinin olmaması hem değerlendirmelerin yararlılığını hem de sonuçların değerlendiriciler için savunulabilirliğini olumsuz etkilemiştir. Değerlendirmelerde, performansı açık ve savunulabilir bir gerekçeye dayalı olarak yorumlamak için iyi tanımlanmış kriterler kullanılmalıdır. Mevcut öğretmen değerlendirme süreçlerinde açık ölçüt veya standartların olmaması da eleştirilmektedir ve birçok çalışma, öğretmen değerlendirme sistemlerine belirli standartların dahil edilmesinin önemini ortaya koymaktadır (Collins, 2009; Donahue, 2016; Ilgaz, 2011; Lillejord vd., 2018; Kimball, 2001; Kraft & Gilmour, 2017; Royal & Tossman, 2009).

Belirlenen ihtiyaçlar doğrultusunda bu öğretmen değerlendirme modelinin amacı “öğretmen niteliklerini değerlendirmek ve değerlendirme sürecine bağlı olarak öğretmen mesleki gelişimine sistematik destek sağlamak” olarak belirlenmiştir. Hesap verebilirliği sağlayan ve öğretmen etkinliğini belirlemede tek kaynak olarak öğrenci gelişiminin ölçülerini kullanan modeller, öğrenenler olarak öğretmenlerin ihtiyaçlarını fark etmekte başarısız olmuşlardır. Öğrencilerin sınav puanları ile yapılan öğretmen değerlendirmelerin kesin sonuçlar ortaya koyduğu düşünülse de bu durum eğitim felsefecilerinin ve araştırmacılarının onlarca yıldır eğitim hakkında bildikleri şeylerle. (Öğretme ve öğrenme doğası gereği belirsiz ve karmaşıktır) doğrudan çelişir (Ford & Hewitt, 2020). Öğretmenin mesleki gelişimi ile öğretimin etkinliğinin artması umulmaktadır. Ancak öğretimi değil, öğretmenin niteliğini artırmak için kullanılan

teşvikleri tercih etmek ve başarısız öğretmeni cezalandırmak, tüm dünyada eğitim reformları için aldatıcı ve hatta ölümcüldür (Fullan, 2011). Öte yandan, Amerika Birleşik Devletleri'nde ve diğer ülkelerde öğretmenlerin başarısının öğrenci test sonuçlarıyla ölçüldüğü "büyüme modelleri" veya "katma değerli modeller" in kullanımı azalmaya devam ederken ve öğretmenlerin mesleki gelişimini destekleyen değerlendirmeler önem kazanmaya başlamıştır (Close ve diğerleri, 2020). Bu durumun, bu çalışma için özel olarak tasarlanmış öğretmen değerlendirme modelinin amacı ile örtüştüğü söylenebilir Ayrıca öğretmen değerlendirme alanında yapılan doktora tezleri ve araştırmalar incelendiğinde bulguların öğretmenin mesleki gelişimini destekleyen değerlendirme modellerinin gerekliliğini ve önemini ortaya koyduğu görülmektedir (Bige, 2014; Evans, 2019; Fowler, 2001; Kimball, 2001; La Masa, 2005; Marzano, 2012; Moss, 2015; Nilsen, 2006; Taylor & Tyler, 2012; Süzen, 2007; Zarro, 2005).

Katılımcılara bu model kapsamındaki değerlendirme sonuçlarının kendi ihtiyaçlarını ortaya çıkaracağı için okuldaki hizmet içi eğitim faaliyetlerine de yön vereceğini belirtmişlerdir. Bu bulgu, TALIS 2018 verileriyle benzerlikler göstermektedir. TALIS 2018 sonuçları, Türkiye'deki öğretmenlerin, öğretmenin ön bilgi ve ihtiyaçları dikkate alınarak gerçekleştirilen hizmet içi eğitimin en etkili olduğuna inandıklarını ortaya koymuştur. Araştırmalar ayrıca öğretmenlerin ihtiyaçlarının doğru belirlendiği okullarda öğretmenlerin profesyonel olarak gelişebileceğini göstermiştir (İçel, 2008; Jiang vd., 2015; Kini ve Podolsky, 2016; Kraft ve Papay, 2014).

Araştırmaya katılanlara belirledikleri amaç doğrultusunda hangi yeterlilik alanlarının değerlendirilmesi ve iyileştirilmesi gerektiği sorusu sorulmuştur. Ortaya çıkan yeterlilik alanlarının (planlama ve hazırlık, öğretim, yansıtıcı düşünme, iletişim ve işbirliği ve mesleki sorumluluklar) önde gelen öğretmen değerlendirme modellerindeki alanlarla örtüştüğü söylenebilir (Clayton, 2017; Danielson; 2013; Marzano ve Toth, 2013; Marzano ve Simms, 2014; Shinkfield ve Stufflebeam, 1997; Sloat ve diğerleri, 2017; Toch ve Rothman, 2008; Bender, 2005; Danielson, 2007; Danielson ve McGreal, 2000; Deiro, 2005; Egan, 2010; Geng ve diğerleri, 2019; Graham ve diğerleri, 2015; Graham ve Berman, 2018).

Araştırmaya katılan öğretmenler, müdürler ve uzmanlar bu okulda yapılan değerlendirmelerin sınıftaki gerçek durumu yansıtmadığını ve sınıfta olup bitenlerin gözlemlenmesi gerektiğini açıkça belirtmişlerdir. Bu nedenle bu modelde temel değerlendirme aracı olarak sınıf gözlemleri kullanılmıştır. Gözlemin öğretimi nasıl etkilediğini araştıran çoğu çalışma, gözlemin öğretimi iyileştirebileceğini veya en azından öğretmenlerin öğretimlerini geliştirmelerine yardımcı olabileceğini ortaya koymuştur (Donaldson ve diğerleri, 2014; Firestone ve Donaldson; Özbek ve Taneri, 2019). Ayrıca bu model kapsamında gözlem formları yardımıyla hem öğretim süreci hem de öğretmenin diğer yeterlikleri ile ilgili birden fazla kaynaktan çoklu veriler toplanması amaçlanmıştır. Değerlendirme sonuçlarına göre doğru kararlar verebilmek için birden fazla kaynaktan veri toplamak ve güvenilirliği sağlamak önemlidir. Bu nedenle bir sınıf öğretmenin etkililiği gözlemlenirken, öğretmenin farklı derslerdeki, farklı konulardaki ve farklı ders bölümlerindeki (yeni bir konunun öğretimi, bir konunun pekiştirilmesi, değerlendirilmesi vb.) performansı birden fazla kişi tarafından değerlendirilmelidir (Marzano, 2012; McQueen, 2022). Birçok araştırma ve literatür kaynağında öğretmenler, mesleki uygulamalarını doğru verilerle gözden geçirmek için çoklu sınıf gözlemleri talep etmektedir (Cohen ve Goldhaber, 2016; Collins, 1999; Ford ve Hewitt, 2020; Fulton, 2019; Ilgaz, 2011; Jiang, Spote ve Luppescu, 2015; Kaplan, 2019; Kimball, 2001; Winslow, 2015; Stronge, 2006; Süzen, 2007; Yılmaz, 2017). Bu çalışma kapsamında tasarlanan değerlendirme modelinde gözlem formlarının ve planlı gözlem sürecinin gözlemden kaynaklanabilecek olumsuz durumların önüne geçeceği tahmin edilmektedir. Geliştirilen gözlem formları, öğretim için kapsamlı ve ayrıntılı ölçütler içermektedir. Bu formların ön pilot çalışmaları yapılmış ve geçerliliği için gerekli düzenlemeler yapılmıştır. Formu kullanacak kişilere uygulamaya dayalı eğitimlerin verilmesinin, değerlendiricilerin yetkinliklerini artıracığı tahmin edilmektedir. Öğretmenlerin birden fazla gözlem yaparak çoklu ve tekrarlı ölçümleri değerlendirecek olması bu model kapsamında geliştirilen formları güçlü kılmaktadır. Model kapsamında, gözlem öncesinde öğretmenlerle yapılan görüşmeler yoluyla gözlemlenecek ders süreci hakkında gerekli bilgiler toplanır ve gözlem sonrasında da ders süreci öğretmenle birlikte değerlendirilir. Böylece

öğretmenin ders süreci boyunca kendini ifade etmesine izin verilmiş ve oluşabilecek yanlış anlamaların önüne geçilmiştir.

Bu araştırmada geliştirilen öğretmen değerlendirme modeli, okulda görev yapan paydaşların (sınıf öğretmenleri, müdürler, uzmanlar) katılımıyla geliştirilmiştir. Öğretmenler, öğretmen değerlendirme modellerinin planlanması ve uygulanmasında söz sahibi olmayı önemsemektedirler. Öğretmenlerin bakış açısıyla geliştirilen değerlendirme süreçleriyle öğretmen çabasını ve etkililiğini nasıl değiştirebileceği hakkında bilinenler nispeten sınırlıdır (Tuytens ve Devos, 2013; Taylor ve Tyler, 2012). Öğretmenlerin öğretmen değerlendirme sistemleri geliştirmeye yönelik bakış açılarını anlamak için sınırlı sayıda çalışma yapılmıştır. Yapılan araştırmalar öğretmenlerin değerlendirme hakkındaki görüş ve önerilerinin alınmasının öğretmenlerin değerlendirmeyi benimsemesini ve uygulamaya istekliliğini arttırdığını ve bu görüşlerin eğitim sistemleri ve politika yapıcılar için faydalı olacağını ortaya koymuştur (Fowler, 2001; Jiang ve diğerleri, 2015). Ayrıca öğretmen değerlendirme sisteminin geliştirilmesinde öğretmenlerin ve diğer paydaşların görüşlerinin dahil edilmesinin ve birlikte geliştirilmesinin önemini vurgulayan araştırma ve literatür çalışmaları, genel olarak öğretmen değerlendirme sistemlerinin daha etkili olacağını göstermiştir (ESSA,2019; McQueen, 2022; Paufler ve diğerleri, 2020). Öğretmen değerlendirme programlarının geliştirilmesine dahil olan öğretmenler, öğrenciler, veliler, yöneticiler ve yetkililer, öğretmen değerlendirme sistemlerinin önemini anlarlar (ESSA, 2019). Öğretmenin bu sistemin etkililiğine ilişkin algısını dikkate alan öğretmen değerlendirme sistemlerinin oluşturulması, bu değerlendirme sisteminin öğrenci başarısı ve öğretmen mesleki gelişimi açısından ne kadar etkili olabileceğini belirlemeye yardımcı olabilir (McQueen, 2022).

Bu araştırma, her okul, öğretmen ve müdür tarafından kullanılabilecek bir öğretmen değerlendirme modeli önermekten ziyade, bir okulun kültürüne ve görev yapan öğretmen ve müdürlerin ihtiyaçlarına yönelik etkili bir öğretmen değerlendirme modelinin nasıl tasarlanacağına rehberlik etmesi bakımından önemlidir. Okullar için geliştirilecek öğretmen değerlendirme modelleri, öğretmenler, müdürler ve okulda görev yapan diğer ilgili kişilerin katılımıyla yapılmalıdır. Aksi takdirde, hali hazırda

geliştirilmiş olan değerlendirme modelleri ne kadar etkili olursa olsun, öğretmenlerin önyargılı tepkileri bu modellerin uygulanmasını engelleyebilir.

Öğretmenler, öğrenci başarı puanları ile değerlendirilmekten memnun değildir, bu değerlendirmeler onlara yaptırımlar getirmektedir, öğretmenler bu değerlendirme modellerine güvenmemekte, geri bildirimleri dikkate almamakta ve bu sürecin adil olmadığını düşünmektedirler. Bu durumda öğretmen değerlendirme modelleri veya sistemleri öğretmenlerin mesleki gelişimini desteklemeli ve etkinleştirmelidir. Öğretmen değerlendirme modelleri, öğretmenlerin önemli niteliklerini ayrıntılı ölçütlerle değerlendirmeyi amaçlamalı ve bu ölçütlerde belirlenen zayıflıkları giderecek mesleki gelişim faaliyetleri planlanmalıdır.

Araştırma bulguları okulun kültürüne, ekosistemine ve paydaşların birbirleriyle olan ilişkilerine özel bir model geliştirmenin gerekliliği ortaya çıkarmıştır. Ülke genelinde performans değerlendirme sistemleri geliştirme ve kullanma eğilimi olmasına rağmen, genel olarak öğretmenin belirli maddelerle puanlanarak değerlendirildiği öğretmen değerlendirme modellerinin özerk ve okula özgü olması gerekmektedir. Bu nedenle ilerideki araştırmalar yardımıyla geliştirilecek olan öğretmen değerlendirme modellerinin okullarda uygulanabilmesi için gerekli esnekliğe sahip olması ve mümkünse okula özgü olması oldukça önemlidir.

N. TEZ İZİN FORMU / THESIS PERMISSION FORM

(Please fill out this form on computer. Double click on the boxes to fill them)

ENSTİTÜ / INSTITUTE

Fen Bilimleri Enstitüsü / Graduate School of Natural and Applied Sciences ☐

Sosyal Bilimler Enstitüsü / Graduate School of Social Sciences ☒

Uygulamalı Matematik Enstitüsü / Graduate School of Applied Mathematics ☐

Enformatik Enstitüsü / Graduate School of Informatics ☐

Deniz Bilimleri Enstitüsü / Graduate School of Marine Sciences ☐

YAZARIN / AUTHOR

Soyadı / Surname : Güneş Savul

Adı / Name : Emel

Bölümü / Department : Eğitim Bilimleri, Eğitim Programları ve Öğretim / Educational Sciences, Curriculum and Instruction

TEZİN ADI / TITLE OF THE THESIS (İngilizce / English): Developing A Participatory Model for Teacher Evaluation and Professional Development: A Case Study

TEZİN TÜRÜ / DEGREE: Yüksek Lisans / Master ☐ Doktora / PhD ☒

1. Tezin tamamı dünya çapında erişime açılacaktır. / Release the entire work immediately for access worldwide. ☐

2. Tez iki yıl süreyle erişime kapalı olacaktır. / Secure the entire work for patent and/or proprietary purposes for a period of two years. * ☐

3. Tez altı ay süreyle erişime kapalı olacaktır. / Secure the entire work for period of six months. * ☒

* Enstitü Yönetim Kurulu kararının basılı kopyası tezle birlikte kütüphaneye teslim edilecektir. / A copy of the decision of the Institute Administrative Committee will be delivered to the library together with the printed thesis.

Yazarın imzası / Signature

Tarih / Date

(Kütüphaneye teslim ettiğiniz tarih. Elle doldurulacaktır.)

(Library submission date. Please fill out by hand.)

Tezin son sayfasıdır. / This is the last page of the thesis/dissertation.