INVESTIGATING PRE-SERVICE AND IN-SERVICE MATHEMATICS TEACHERS' ASSESSMENT LITERACY: SUGGESTIONS FOR AN EFFECTIVE TEACHER EDUCATION PROGRAM

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

BEYZA ŞAHİN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
MATHEMATICS AND SCIENCE EDUCATION

SEPTEMBER 2023

Approval of the thesis:

INVESTIGATING PRE-SERVICE AND IN-SERVICE MATHEMATICS TEACHERS' ASSESSMENT LITERACY: SUGGESTIONS FOR AN EFFECTIVE TEACHER EDUCATION PROGRAM

submitted by **BEYZA ŞAHİN** in partial fulfillment of the requirements for the degree of **Master of Science in Mathematics and Science Education Department, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences** _____

Prof. Dr. Mine Işıksal Bostan
Head of Department, **Mathematics and Science Education** _____

Assoc. Prof. Dr. Bülent Çetinkaya
Supervisor, **Department name, Institution Name** _____

Examining Committee Members:

Prof. Dr. Ayhan Kürşat Erbaş
Mathematics and Science Education, METU _____

Assoc. Prof. Dr. Bülent Çetinkaya
Mathematics and Science Education, METU _____

Assoc. Prof. Dr. Nazan Sezen Yüksel
Mathematics and Science Education, HÜ _____

Date: 06.09.2023

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, Surname: BEYZA ŞAHİN

Signature :

ABSTRACT

INVESTIGATING PRE-SERVICE AND IN-SERVICE MATHEMATICS TEACHERS' ASSESSMENT LITERACY: SUGGESTIONS FOR AN EFFECTIVE TEACHER EDUCATION PROGRAM

ŞAHİN, BEYZA

M.S., Department of Mathematics and Science Education

Supervisor: Assoc. Prof. Dr. Bülent Çetinkaya

September 2023, 191 pages

This study has three primary objectives: assessing the assessment literacy of pre-service and in-service middle school mathematics teachers, exploring their perceptions of the areas that need improvement in undergraduate assessment education, and examining their recommendations for improving assessment literacy in teacher education programs. Both quantitative and qualitative research designs were employed in the study. The "Assessment Literacy Inventory" was used to collect data from 34 pre-service and 41 in-service middle school mathematics teachers at the end of the 2021-2022 spring semester. Additionally, semi-structured interviews were conducted with five pre-service and five in-service teachers, purposively selected from voluntary participants. Findings revealed low assessment literacy among middle school pre-service and in-service teachers. Also, according to inferential statistical results, there was no significant difference in the assessment literacy scores between pre-service and in-service teachers. Pre-service teachers scored (out of 5) highest in administering, scoring, and interpreting the assessment results (M=3.35, SD= 1.01) and lowest in creating accurate techniques for grading (M=1.85, SD= 1.01), while in-service teachers scored highest in selecting appropriate assessment methods

(*M*=3.32, *SD*= 1.06) and lowest in creating accurate techniques for grading (*M*=1.85, *SD*= 0.88). Moreover, participants identified various areas that need improvement regarding both theoretical and practical aspects of assessment education in the teacher education program. Finally, participants' recommendations for improving assessment education were grouped into these categories: applying assessment skills practically, restructuring assessment courses, aligning course content with daily assessment challenges, promoting collaboration, encouraging active participation, providing more learning opportunities, and highlighting assessment's importance in teacher education programs.

Keywords: middle school pre-service mathematics teachers, middle school in-service mathematics teachers, assessment, assessment literacy, teacher education program

ÖZ

HİZMET ÖNCESİ VE HİZMET İÇİ MATEMATİK ÖĞRETMENLERİNİN ÖLÇME-DEĞERLENDİRME OKURYAZARLIĞININ İNCELENMESİ: ETKİLİ BİR ÖĞRETMEN EĞİTİM PROGRAMI İÇİN ÖNERİLER

ŞAHİN, BEYZA
Yüksek Lisans, Matematik ve Fen Bilimleri Eğitimi Bölümü
Tez Yöneticisi: Doç. Dr. Bülent Çetinkaya

Eylül 2023 , 191 sayfa

Bu çalışmanın üç temel amacı vardır: hizmet öncesi ve hizmet içi ortaokul matematik öğretmenlerinin ölçme-değerlendirme okuryazarlığını değerlendirmek, lisansüstü değerlendirme eğitiminde iyileştirilmesi gereken alanları keşfetmek ve öğretmen eğitim programlarında ölçme-değerlendirme okuryazarlığını geliştirmek için önerilerini incelemektir. Çalışmada hem nicel hem de nitel araştırma desenleri kullanılmıştır. Veri toplama aracı olarak "Ölçme-Değerlendirme Okuryazarlığı Envanteri" kullanılmış ve bu envanterle veriler, 2021-2022 bahar döneminin sonunda 34 öğretmen adayı ve 41 hizmet içi ortaokul matematik öğretmeninden toplanmıştır. Ayrıca, gönüllü katılımcılardan amaçlı olarak seçilen beş hizmet öncesi ve beş hizmet içi öğretmenle yarı yapılandırılmış görüşmeler gerçekleştirilmiştir. Bulgular, ortaokul hizmet öncesi ve hizmet içi öğretmenler arasında ölçme-değerlendirme okuryazarlığının düşük olduğunu göstermektedir. Çıkarımsal istatistik sonuçlarına göre, hizmet öncesi ve hizmet içi ortaokul matematik öğretmenlerinin ölçme-değerlendirme okuryazarlığı skorları arasında anlamlı

bir fark bulunmamıştır. Hizmet öncesi öğretmenler, değerlendirme sonuçlarını uygulama, puanlama ve yorumlama konusunda 5 üzerinden en yüksek puanı alırken (*M*=3.35, *SD*= 1.01), en düşük puanı doğru derecelendirme teknikleri oluşturma konusunda almışlardır (*M*=1.85, *SD*= 1.01). Hizmet içi öğretmenler ise uygun değerlendirme yöntemlerini seçme konusunda en yüksek puanı alırken (*M*=3.32, *SD*= 1.06), en düşük puanı doğru derecelendirme teknikleri oluşturma konusunda almışlardır (*M*=1.85, *SD*= 0.88). Ayrıca, katılımcılar öğretmen eğitim programındaki değerlendirme eğitimi konusunda hem teorik hem de pratik yönlerde iyileştirilmesi gereken çeşitli alanları belirlemişlerdir. Katılımcıların değerlendirme eğitimini geliştirmeye yönelik önerileri, değerlendirme bilgi ve becerilerinin pratiği, derslerin yapısı ve sıralaması, ders içeriğinin günlük uygulamalarla uyumlu hale getirilmesi, işbirliğine dayalı öğrenme topluluğunun teşvik edilmesi, aktif katılımın teşviki, daha fazla öğrenme fırsatının sunulması ve öğretmen yetiştirme programlarında değerlendirmenin önemi gibi kategorilere ayrılmıştır.

Anahtar Kelimeler: hizmet öncesi ortaokul matematik öğretmen adayları, hizmet içi ortaokul matematik öğretmen adayları, değerlendirme, ölçme-değerlendirme okuryazarlığı, öğretmen yetiştirme programı

To my family and all loved ones

ACKNOWLEDGMENTS

I would like to express my heartfelt gratitude by extending my thanks to my supervisor, Assoc. Prof. Dr. Bülent Çetinkaya, for his unwavering support, encouragement, and guidance throughout my academic journey. He provided invaluable advice and assistance at every phase of the research. Each meeting with him broadened my horizons, both in terms of academic knowledge and personal growth. Collaborating with him has been a privilege, and I am truly thankful for his patience and continuous support during this study. Thanks a million.

I also wish to express my appreciation to my committee members, Prof. Dr. Ayhan Kürşat ERBAŞ, and Assoc. Prof. Dr. Nazan Sezen YÜKSEL, for their valuable suggestions, contributions, and feedback.

I would like to express my gratitude to my loving family who has been with me at every stage of my thesis work: my mother, Nurgül ŞAHİN; my father, İrfan ŞAHİN; and my dear sisters, Büşra and Kübra ŞAHİN. I want to convey that this journey would have been much more challenging without your support, encouragement, and boundless love. You are not only my family but also a source of inspiration for me. Thank you for always being there for me.

I would like to thank Sena KORKMAZ, who was one of the biggest supporters throughout my thesis work. She was always there for me and always believed in me. Her understanding, motivational words and contributions to my thesis are invaluable to me. Thank you for giving me the courage I needed to take the steps toward my life.

Moreover, I would like thank to my friends Behiye Nur KESGİN, Bengisu ÖZDEN, Botan KAYA, Deniz NUR, Ender DOĞRU, Mahmut KILIÇ, and Şaziye Nur ERTÜRK. I appreciate your unwavering support during this demanding journey. I consider myself fortunate to have friends as supportive as you.

TABLE OF CONTENTS

ABSTRACT v

ÖZ vii

ACKNOWLEDGMENTS x

TABLE OF CONTENTS xi

LIST OF TABLES xvii

LIST OF ABBREVIATIONS xx

CHAPTERS

1 INTRODUCTION 1

1.1 Purpose of the Study 3

1.2 Research Questions 4

1.3 Significance of the Study 5

1.4 Definition of the Important Terms 8

2 LITERATURE REVIEW 11

2.1 Assessment 11

2.1.1 Purposes of Assessment 15

2.1.2 Standards of Assessment 17

2.1.3 The Process of Assessment 21

2.1.3.1 Generating or Choosing the Assessment 22
2.1.3.2 Administering the Assessment 23
2.1.3.3 Scoring the Assessment Results 24
2.1.3.4 Interpreting and Using the Assessment Results 25
2.1.3.5 Communicating the Assessment Results 26
2.2 Assessment Literacy 27
2.2.1 Assessment Literacy and Teacher Education Programs 30
2.2.2 Studies on Assessment Literacy 33
2.2.2.1 Developing Assessment Literacy Inventory 33
2.2.2.2 Pre-service and In-service Teacher's Assessment Literacy 36
2.3 Summary of the Literature Review 44
3 METHODOLOGY 47
3.1 Design of the Study 47
3.2 Context of the Study 49
3.3 Participants of the Study 51
3.3.1 First Stage of the Sample Selection 52
3.3.2 Second Stage of Sample Selection 54
3.4 Data Collection Tools 57
3.4.1 Assessment Literacy Inventory 58
3.4.2 Interview Protocol 62
3.5 Data Collection Process 64
3.6 Data Analysis 65

3.6.1 Analysis of Data from the Assessment Literacy Inventory . . . 66

3.6.2 Analysis of Data from Interview Transcripts 66

3.7 The Trustworthiness of the Study 70

3.8 Ethical Considerations of the Study 72

3.9 Assumptions of the Study 73

4 FINDINGS 75

4.1 Descriptive Statistics 75

4.2 Inferential Statistics 77

4.3 Participants' Views About the Shortcomings of the Teacher Education Program Regarding Assessment Literacy 79

4.3.1 General Deficiencies 79

4.3.1.1 General Deficiencies in Theoretical Knowledge 81

4.3.1.2 General Deficiencies in Practice/experience 82

4.3.2 Deficiencies in the Selecting and Generating Appropriate Assessment Methods 85

4.3.2.1 Theoretical Deficiencies 86

4.3.2.2 Practice/experience Deficiencies 88

4.3.3 Deficiencies in the Administering, Scoring, and Interpreting the Assessment Results 88

4.3.3.1 Theoretical Deficiencies 90

4.3.3.2 Practice/experience Deficiencies 92

4.3.4 Deficiencies in the Using Assessment Results for Making Educational Decisions 93

4.3.4.1 Theoretical Deficiencies 93

4.3.4.2 Practice/experience Deficiencies 95

4.3.5 Deficiencies in the Creating Accurate Techniques for Grading . 95

4.3.5.1 Theoretical Deficiencies 96

4.3.6 Deficiencies in the Explaining Assessment Findings to Students, Colleagues, and Parents 97

4.3.6.1 Theoretical Deficiencies 98

4.3.6.2 Practice/experience Deficiencies 100

4.4 Suggestions for Enhancing Assessment Literacy in a Teacher Education Program 100

4.4.1 What: Content of a Program to Improve Assessment Literacy 101

4.4.1.1 Practical Application of Assessment Knowledge and Skills 101

4.4.1.2 The Structure and Sequencing of Assessment Related Courses in The Teacher Education Program 104

4.4.1.3 Enriching The Content of Assessment Courses in Line with Assessment Challenges and Daily Assessment Practices 107

4.4.2 How: Method to Improve Assessment Literacy 110

4.4.2.1 Building a Collaborative Learning Community 111

4.4.2.2 Advocating Active Engagement 114

4.4.2.3 Providing More Learning Opportunities Regarding Assessment 117

4.4.3 Why: Orientation to Improve Assessment Literacy 118

4.4.3.1 Cultivating Reflective Practitioner 118

5 DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS 121

5.1 Assessment Literacy Levels of Pre-service and In-service Teachers . . 121

5.2 Deficiencies of the Teacher Education Program in Developing Assessment Literacy 124

5.2.1 Deficiencies of Theoretical Knowledge 125

5.2.2 Deficiencies in Practice/Experience 129

5.2.3 Gaps Between Theory and Practice 134

5.3 Suggestions of the Teacher Education Program in Developing Assessment Literacy 136

5.3.1 What: Content of a Program to Improve Assessment Literacy . 136

5.3.1.1 Practical Application of Assessment Knowledge and Skills 136

5.3.1.2 The Structure and Sequencing of Assessment Related Courses in the Teacher Education Program 138

5.3.1.3 Enriching The Content of Assessment Courses in Line with Assessment Challenges and Daily Assessment Practices 139

5.3.2 How: Method to Improve Assessment Literacy 142

5.3.2.1 Building a Collaborative Learning Community 142

5.3.2.2 Advocating Active Engagement 144

5.3.2.3 Providing More Learning Opportunities Regarding Assessment 146

5.3.3 Why: Orientation to Improve Assessment Literacy 147

5.3.3.1 Cultivating Reflective Practitioner 147

5.4 Conclusions 148

5.5 Implications for Educational Practices 149

5.6 Recommendations for Further Studies 152

5.7 Limitations of the Study 153

REFERENCES 155

APPENDICES 174

A MSE 305: ASSESSMENT OF LEARNING IN SCIENCE AND MATHEMATICS COURSE SYLLABUS (FALL 2022-2023) 175

B ASSESSMENT LITERACY INVENTORY 177

C INTERVIEW PROTOCOL 185

D NORMALITY HISTOGRAMS 187

E CONSENT FORM 189

F METU HUMAN SUBJECTS ETHICS COMMITTEE APPROVAL 191

LIST OF TABLES

TABLES

Table 3.1 Major Characteristics of the Participants Who Participated in the First Stage of the Study 53

Table 3.2 Characteristics of Pre-service Teachers Who Participated in the Interviews 55

Table 3.3 Characteristics of In-service Teachers Who Participated in the Interviews 57

Table 3.4 Competence Areas and the Corresponding Items for the Adapted Version of the Assessment Literacy Inventory (Bütüner et al., 2010) 60

Table 3.5 Comparison of the Assessment Competency Areas in the Original Inventory and the Competencies Stipulated by the Ministry of National Education in Turkey (Bütüner et al., 2010) 61

Table 4.1 Descriptive Statistics Results of Pre-service and In-service Teachers' Assessment Literacy Levels 76

Table 4.2 Results of Skewness and Kurtosis Values 78

Table 4.3 Tests of Normality by The Shapiro-Wilk Test Statistics 78

Table 4.4 General Deficiencies of the Teacher Education Program Regarding Assessment Literacy 80

Table 4.5 Deficiencies of the Teacher Education Program in Selecting and Generating Proper Assessment Methods 86

Table 4.6 Deficiencies of the Teacher Education Program in Administering, Scoring, and Interpreting the Assessment Results 90

Table 4.7 Deficiencies of the Teacher Education Program in Using Assessment
Results for Making Educational Decisions 93

Table 4.8 Deficiencies of the Teacher Education Program in Creating Accurate Techniques for Grading 96

Table 4.9 Deficiencies of the Teacher Education Program in Explaining Assessment Findings to Students, Colleagues, and Parents 98

Table 4.10 Suggestions For Enhancing Assessment Literacy in the Content of Teacher Education Program 102

Table 4.11 Suggestions For Enhancing Assessment Literacy in the Method of Teacher Education Program 111

Table 4.12 Suggestions For Enhancing Assessment Literacy in the Orientation of Teacher Education Program 118

LIST OF ABBREVIATIONS

AFT	American Federation of Teachers
EME	Elementary Mathematics Education
MoNE	Ministry of National Education
NCTM	National Council of Teachers of Mathematics
NCME	National Council on Measurement in Education
NEA	National Education Association

CHAPTER 1

INTRODUCTION

In this century, it has become an important skill for all educators to have the knowledge of both how to teach and how to assess what they have taught, and the ability to make decisions based on these assessments (Huang & He, 2016). This has made measurement and assessment an integral and complementary part of the educational process and instructional programs (Pilcher, 2001). This critical relationship between assessment and learning is essential for higher-quality education (Baird et al., 2017). According to Brookhart (1999), "The quality of instruction in any classroom turns on the quality of the assessments used there" (p. 20). Integrating appropriate assessment into teaching and using the data obtained as a result of assessment for teaching purposes significantly increases student success (Mertler & Campbell, 2005; Stiggins , 2005). With the assessment, teachers determine the readiness and learning of their students, interpret their learning outcomes, and make up for the students' deficiencies with the results obtained, thus increasing the learning quality of the students (Black & Wiliam, 1998). Assessment contributes significantly to achieving learning objectives throughout the instructional process, not just at the end of teaching. It is essential to incorporate assessment at every teaching stage to ensure effective learning outcomes (Wiliam, 2007). However, to achieve this successfully, teachers must possess sufficient knowledge and skills in assessment. Teachers' awareness and assessment competence are vital for effective instruction (Looney et al., 2018). For this reason, it is one of the indispensable elements of a quality teaching-learning process for a qualified teacher to know both how to teach and how to assess what he/she teaches (Karaman & Şahin, 2014).

Teachers who possess solid knowledge and adequate skills in assessment are

referred to as assessment-literate. Popham (2014) defined assessment literacy as the understanding and familiarity of educators with essential assessment concepts and procedures that are likely to influence educational judgments. Stiggins (1995) stated that teachers who are literate in assessment should have sufficient knowledge, skills, and experience on what to assess, how to evaluate best what the student has learned, what factors will negatively affect assessment, and how to prevent them. Teachers need to become assessment-literate individuals to enhance the quality of teaching (Mertler & Campbell, 2005). Therefore, it is essential for teacher candidates who receive training from teacher education programs to enter their professional lives as individuals who are knowledgeable in assessment (Zolfaghari & Ahmadi, 2016). At this point, education faculties have a great responsibility to prepare future teachers for educational assessment (Lai & Waltman, 2008). With the importance of educating teachers who are assessment literate, some measurement tools have been developed to scale the assessment literacy of teachers and teacher candidates (DeLuca et al., 2016; Lai & Waltman, 2008; Mertler & Campbell, 2005; Plake et al., 1993). However, studies in the related literature reported that the assessment literacy of both teachers and teacher candidates is low and should be improved (Brookhart, 2001; DeLuca & Klinger, 2010; Mertler & Campbell, 2005; Plake et al., 1993; Popham, 2004; Stiggins , 1995; Volante & Fazio, 2007). According to research studies, teachers are not ready or able to integrate assessment effectively into their teaching and especially new teachers have low self-confidence in assessment (Mertler, 2004; Ogan-Bekiroglu & Suzuk, 2014; Volante & Fazio, 2007). Popham (2018) attributed the inadequate assessment literacy of in-service teachers to their failure to utilize the concept and procedures of assessment when making instructional decisions. Stiggins (2004) argued that teachers always use the same ways to evaluate students because they are not given a chance to learn new understandings and practices about assessment. In addition, K. Koh and Velayutham (2009) stated that teachers tend to use traditional assessment methods and avoid challenging assessment tasks. The researchers noted that teachers need professional help and support in designing more effective assessment tools. Moreover, Bol and Strage (1996) expressed that teachers' teaching objectives and practices are not compatible. They tend to demand low cognitive processing for assessment and lack in criticizing their assessment tools. Similar to in-service teachers, studies have also shown that pre-service teachers

also were unable to effectively incorporate assessment literacy into their assessment practices (Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk, 2014; Siegel & Wissehr, 2011; Yilmaz-Tuzun, 2008). It has been revealed that teacher candidates lack sufficient knowledge and skills regarding various assessment methods, and they do not have the opportunity to question these methods, receiving limited experience in assessment training (Volante & Fazio, 2007). DeLuca et al. (2013) stated that pre-service teachers in teacher education programs could not acquire sufficient assessment knowledge and skills due to the inadequacy of pre-service training in terms of assessment. In his study, Mertler (1999) revealed that teacher candidates, despite taking assessment courses, were unable to acquire the necessary skills in student assessment even upon completion of the course. The results of the aforementioned studies show the need to develop and improve teacher education programs to enhance the assessment literacy levels of future teachers.

Despite the critical importance of assessment literacy for better teaching, there is a lack of research dedicated to investigation of the assessment literacy of teachers and teacher candidates and how their assessment literacy can be improved from their perspectives (Bijsterbosch et al., 2019; DeLuca et al., 2013; Graham, 2005; Mertler, 2004; Ogan-Bekiroglu & Suzuk, 2014; Volante & Fazio, 2007; Ye, 2023). Thus, it is essential to identify pre-service and in-service teachers' needs and expectations of a teacher education program that can better support their assessment literacy.

1.1 Purpose of the Study

As mentioned above, although teachers need to be assessment literate to improve students' learning, assessment literacy levels of teachers and teacher candidates are low (DeLuca et al., 2013; Graham, 2005; Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk, 2014; Popham, 2014; Stiggins , 2004). Stiggins (2004) declared, "Few teachers are prepared to face the challenges of classroom assessment because they have not been given the opportunity to learn to do so" (p. 762). Volante and Fazio (2007) noted that there is a lack of research into teachers' assessment literacy to understand it. Moreover, the researchers argue that research for assessment literacy should begin in faculties of education since they prepare adequately for

the measurement and assessment of prospective teachers. According to these, the study has three primary purposes. The first aim of this study is to investigate the assessment literacy level of middle school mathematics pre-service and in-service teachers. The second aim is to investigate the deficiencies in the assessment training received by both pre-service and in-service mathematics teachers during their undergraduate education. The final aim is to gather suggestions from pre-service and in-service mathematics teachers on how the teacher education program can be more effective in enhancing assessment literacy and improving the quality of assessment in undergraduate education.

1.2 Research Questions

In light of the three purposes presented above, the following research questions guided this thesis:

1. What is the assessment literacy level of pre-service and in-service middle school mathematics teachers?

   (a) Is there a significant difference between the assessment literacy scores of pre-service and in-service middle school mathematics teachers?
   The Null Hypothesis: There is no significant difference between the pre-service and in-service teachers' mean assessment literacy inventory scores.

2. What are the perceptions of pre-service and in-service mathematics teachers about the areas that need improvement in undergraduate assessment education that can support the development of their assessment literacy?

3. What do pre-service and in-service mathematics teachers think about how the teacher education program can more effectively develop their assessment literacy?

1.3 Significance of the Study

Teachers and students co-create the learning experience through their actions in the classroom. Assessment is one of the competencies teachers need to possess to support students' learning and help them become proficient learners (Ayalon & Wilkie, 2020; Black & Wiliam, 1998; Volante & Fazio, 2007). While grading is one aspect of assessment, it encompasses much more. Assessment plays a crucial role in gathering information about the attainment of curriculum objectives, the effectiveness of learning-teaching activities, methods, and techniques, and how students organize their knowledge (Suurtamm et al., 2016). In today's education landscape, the importance of assessment has been on the rise, and along with it, the significance of assessment literacy. According to Popham (2018), individuals with the skills to comprehend and apply fundamental assessment concepts and procedures that impact student learning are regarded as assessment literate. Understanding the objective of assessments, being familiar with the different assessment kinds, and applying them effectively are all aspects of assessment literacy for teachers. It entails giving helpful feedback, ensuring valid and reliable assessments, exhibiting cultural sensitivity, and remaining dedicated to professional development (Bijsterbosch et al., 2019). In essence, it is about giving teachers the tools to efficiently assess students' learning and foster their development in the classroom. Furthermore, Popham argues that teachers also need to be assessment literate, yet many teachers lack knowledge of basic assessment concepts and procedures. Assessment literacy allows teachers to make well-informed choices regarding assessment methods, ultimately resulting in enhanced student learning results (Baird et al., 2017). Consequently, it constitutes a pivotal element of effective teaching and should be continually cultivated and honed throughout a teacher's professional journey (S. Brown, 2014; Looney et al., 2018). Despite assessment literacy being a crucial skill for enhancing the quality of teaching and learning, many teachers have low self-confidence in this area. They often struggle to effectively integrate assessment into their practice (Mertler, 2004; Volante & Fazio, 2007). According to the literature, one of the main reasons for the low assessment literacy levels among teachers and teacher candidates is the inadequate assessment education provided by teacher education programs (DeLuca et al., 2013; DeLuca & Klinger, 2010; Graham, 2005; Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk,

2014; Popham, 2014; Volante & Fazio, 2007). Moreover, studies indicated the need for the enhancement of assessment literacy education offered within teacher training programs (Coombe et al., 2020; Ogan-Bekiroglu & Suzuk, 2014; Pastore & Andrade, 2019; Popham, 2009; Volante & Fazio, 2007). Although there have been some studies on assessment literacy (DeLuca & Klinger, 2010; Graham, 2005; Mertler, 2004; Siegel & Wissehr, 2011; Volante & Fazio, 2007), the number of studies conducted for the development of assessment literacy in our country is quite limited (Çambay, 2016; Karaman, 2014; Ogan-Bekiroglu & Suzuk, 2014). Hence, conducting comprehensive research on assessment literacy within teacher education programs, identifying the areas that need improvement, and offering suggestions to address the shortcomings of assessment education will contribute to the revisions and refinements effort of developing pre-service and in-service teachers' assessment literacy.

Several research studies on assessment literacy took into account the document, "Standards for Teacher Competence in Educational Assessment of Students," designed by the National Council on Measurement in Education (NCME), the National Education Association (NEA), and the American Federation of Teachers (AFT) in 1990. This document proposed that teachers should be skilled in i) choosing assessment methods appropriate for instructional decisions, ii) developing assessment methods appropriate for instructional decisions, iii) scoring, and interpreting the results of both externally produced and teacher-produced assessment methods, iv) using assessment results when making decisions about individual students, planning teaching, developing curriculum, and promoting school improvement, v) developing valid student grading procedures that use student assessments, vi) communicating assessment results to students, parents, other lay audiences, and other educators, an recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information. A number of assessment literacy inventories have been developed using these standards. In the studies, although a measurement tool prepared in accordance with the assessment competency domains included in these assessment standards were used, most of the studies were quantitative, and these competency domains have not been thoroughly examined individually (Akayuure, 2021; Gül, 2011; McGee & Colby, 2014; Mertler, 2004; Mertler & Campbell, 2005; Tünkler, 2019). In addition, there are very few studies on teacher

education programs' shortcomings in assessment literacy and suggestions aimed to overcome them. In this study, the assessment literacy levels of middle school pre-service and in-service mathematics teachers were measured according to their assessment competency domains. Additionally, interviews were conducted with the participants to explore the areas that need improvement in their undergraduate education and the recommendations they offered for improving the quality of undergraduate education for each competency area. Thus, the researcher aimed to gain a comprehensive understanding and enriched perspective on the shortcomings of and recommendations for teacher education programs by drawing on the participants' assessment experiences during their undergraduate education. In addition, there are very few studies on teacher education programs' shortcomings in assessment literacy and suggestions aimed to overcome them. Therefore, this study bridges the gap in the literature and varies from other studies in this regard.

Furthermore, national and international studies on assessment literacy have mostly been conducted with teacher candidates (Çambay, 2016). Teacher candidates are individuals who have not yet completed their education, while teachers have finished the teacher education program and gained experience in the classroom. Therefore, there may be differences in the experiences and perspectives of pre-service and in-service teachers regarding assessment education in undergraduate education. Consequently, a study that includes both groups can facilitate the emergence of different perspectives and contribute to a more comprehensive understanding. This study was conducted with middle school mathematics teacher candidates and teachers who have graduated from the same university and have taken an assessment course. Assuming they have similar backgrounds in assessment within the teacher education program, the aim was to provide a profound viewpoint on the deficiencies and suggestions regarding assessment education. This profound viewpoint is expected to be shaped by the assessment experiences that teacher candidates have had more recently during their undergraduate education and the teachers' teaching experiences in their professional fields. Therefore, obtaining the opinions of pre-service and in-service teachers regarding the aspects that need improvement in the assessment education provided in teacher education programs, and providing recommendations for teacher education programs based on these opinions would make a significant

contribution to the relevant literature. Moreover, investigating suggestions for improving assessment courses at the university where the study is carried out will shed light on the development of assessment curricula of universities that offer similar assessment courses.

1.4 Definition of the Important Terms

Assessment: A systematic process of collecting and documenting data on students' knowledge, skills, and dispositions to learning and using that data to enhance and develop teaching (National Council of Teachers of Mathematics, 1995).

Assessment Literacy: It refers to an individual's understanding of assessment approaches needed to achieve various educational goals, the ability to select appropriate assessment methods, the competence to implement them in the classroom, and the capacity to make necessary improvements based on assessment results Mertler (2004).

Pre-service Middle School Mathematics Teachers: A prospective teacher who is enrolled in a teacher education program at a university's education faculty and has undergone the necessary academic training for a specified period before commencing teaching (Chand et al., 2022). The pre-service teachers participating in the study are studying the Elementary Mathematics Education (EME) Program at the university where the research was conducted. These participants are in their final years of a four-year teacher education (3rd and 4th grade) program. They are being prepared to teach middle school students in grades 5 through 8. At the time of the study, all of them had finished the assessment and statistics courses provided by the Department of Mathematics and Science Education.

In-service Middle School Mathematics Teachers: Teachers who are currently engaged in teaching (Ayvaz-Tuncel & Çobanoğlu, 2018). The participating in-service teachers have also graduated from the same department as the pre-service teachers, namely the Elementary Mathematics Education (EME) Program, which is a four-year teacher education program. They are currently teaching in middle school grades 5 to 8. All of the graduated teachers have completed the assessment, statistics, and

teaching experience courses.

Lesson Plan: A lesson plan serves as a structured documentation and guide for teachers, describing their ideas regarding the content to be covered in a lesson and the effective facilitation of students' learning within the designated class duration (Richards & Renandya, 2002). Typically, a good lesson plan is anticipated to consist of three essential elements: learning goals, learning activities, and assessment methods to evaluate students' comprehension ("Lesson Planning | Centre for Teaching Excellence", n.d.).

CHAPTER 2

LITERATURE REVIEW

This study aims to investigate the perceptions of pre-service and in-service middle school mathematics teachers about how the mathematics education they receive at university affects their assessment literacy. Therefore, the study has one main focuses: assessment literacy. In consideration of these research fields, this section presents studies on assessment in general, assessment literacy, the impact of teacher education programs on assessment, and assessment literacy of pre-service and in-service teachers. The first part includes the purposes, standards, and process of assessment sections. The second part will explore assessment literacy, teacher education programs, and studies on assessment literacy in literature.

2.1 Assessment

Lifelong learning is important in our age, and the development of learning skills has become increasingly significant. The education system is student-centered and aims to achieve qualified learning outcomes. Therefore, elements such as instruction, learning, and assessment affect the quality of the education system (Demirtaşlı, 2014). Assessment refers to a systematic process that aims to evaluate students' proficiency levels in key educational domains deemed essential (Popham, 2004). Nitko and Brookhart (2007) define assessment as a systematic gathering of data and input to make specific educational judgments about pupils, programs, curriculum, academies, and policies. Assessment is a form of evaluation that can be verified based on particular predefined objectives or aims, enabling either relative or numerical measures. (Scriven, 1967). Within the context of mathematics education, (National

Council of Teachers of Mathematics, 1995, p. 3) describes assessment as "the process of gathering evidence about a student's knowledge of, ability to use, and dispositions toward, mathematics and making inferences from that evidence for a variety of purposes". By analyzing the assessment results, teachers can decide how much progress has been made toward the targeted behaviors, identify gaps in student learning, and determine whether adequate learning has been achieved (Yılmaz, 2002). Moreover, assessment enables teachers to reconsider their goals and target behaviors, and plan future lessons with these in mind (Popham, 2014). Therefore, an assessment that improves teaching, which makes it more effective and high-quality, is an crucial part of education (Suurtamm et al., 2016). By measuring and evaluating student achievements throughout the learning process, teachers can identify and address variables that impede learning, such as instructional techniques, learning environments, or course materials (Yılmaz, 2002). Moreover, regular and high-quality assessment improves teaching quality by providing teachers with feedback on their instruction and students' progress (Maki, 2002; National Research Council, 2001; Rockman, 2002).

Numerous studies have demonstrated the positive impacts of assessment on student learning (Boud, 2015; T. Guskey, 2003; Irons & Elkington, 2021), highlighting the essential role of assessment in achieving effective teaching and learning (National Council of Teachers of Mathematics, 2014; Turgut & Baykul, 2010). Martinez and Martinez (1992) conducted a study on the effects of repetitive testing and teacher impact on students' performance in an algebra lecture. The study involved 120 students who were divided into four groups using an experimental method, with two groups receiving one test per lecture and two groups receiving three tests per chapter. The study found that continuous testing positively affected learning, and students who had more frequent testing benefited significantly. Black and Wiliam (1998) conducted a study reviewing the literature on formative assessment used by teachers. The researchers found evidence that using techniques that provide regular feedback during the learning process significantly improves and enhances student learning. Thus, they highlighted the crucial role of formative assessment in the classroom, emphasizing that it is essential for teachers to provide ongoing feedback to students on their learning progress. Biggs (1998) conducted research to investigate the position

and significance of summative assessment in the process of learning. He concluded that, similar to formative assessment, summative assessment plays a vital part in enhancing learning. He proposed the integration of the potent relationship between formative and summative assessments to create a comprehensive synthesis. Both kinds of assessments are valuable for learning because they offer different feedback types that can enhance student performance and advancement. While summative assessment gauges student achievement at a specific time, formative assessment provides continuous feedback throughout the learning process. By using both forms of assessment, instructors can better tailor instruction and improve student learning outcomes. More recently, Wiliam (2011) examined various studies on assessment, investigating the positive impact of assessment on learning. The research conducted by the researcher suggests that combining assessment with instruction can increase student attention and advance learning outcomes. The researcher emphasized that for the assessment to positively affect learning, the feedback given as a result of the assessments should be transferred to the students more effectively, and a classroom environment should be created for this. It is vital to use appropriate assessment methods, techniques, and principles at every stage of the learning process for the improvement of teaching (Suurtamm et al., 2016), and one of the greatest responsibilities of teachers is to constantly make judgments and make decisions about the most appropriate teaching and assessment methods for their students (Anderson, 2003; Biggs, 1998; Black & Wiliam, 1998; Hogan, 2007; Popham, 2014).

Thomas et al. (2004) proposed that the educational process comprises three essential components: the curriculum, the instruction, and the assessment. Without any of these elements or if they are not adequately addressed, the entire process is compromised. With a similar approach, Özçelik (1998) argued that for the curriculum-teaching-evaluation triangle to work effectively, it is essential to decide whether applying the teaching methods is effective according to objective judgments from the assessment. The researcher claimed that teaching methods could be improved with correct assessment. Likewise, Rust (2002a) stated that teachers could evaluate the extent of student's learning and their approach to the subject matter and identify which aspects of the lesson interest them through assessment. The components that can provide these and that the assessment must have to be of high

quality and effective (Brown et al., 2013; Reynolds et al., 2008; Young, 1999) are as follows:

- Assessment should be based on a comprehension of how students learn.
- Assessment should be consistent, valid, and reliable.
- The purpose of the assessment should be clarified.
- The criteria of assessment should be intelligible and transparent.
- The amount of assessment applied should be manageable.
- Since students may have different unique distinctions, the assessment should fit in with the needs of these distinctions.
- Assessment techniques should authorize students to take feedback on their learning and performance.
- Assessment should provide chances for self-assessment by allowing students to review what they have learned and teachers to review what they have taught.
- Assessment should be done regularly and a necessary part of the curriculum and lesson plans.

Considering the above components that a good assessment should have, educators have a significant role in assessment (Reynolds et al., 2008; Suurtamm et al., 2016; Thomas et al., 2004; Yılmaz, 2002). Rust (2002b) argued that educators should carefully consider their targeted learning outcomes and plan their assessment stages accordingly. In other words, the researcher stated that educators should allocate time to the assessment procedures and identify learning outcomes to ensure effective teaching and learning. At this point, it is crucial to identify the purposes of the assessment. According to Barnes et al. (2017), assessment tools must be appropriate for the set purposes to improve teaching and student learning. The purposes of assessment are mentioned in the following part.

2.1.1 Purposes of Assessment

Teachers collect data and provide feedback on students' learning, achievement, and developments in teaching practice through assessment (National Research Council, 2001) by applying specific tests or observations to understand what students have learned and what they have not learned (Baird et al., 2017). Although educational institutions make decisions on assessment-related issues, teachers and students are the key drivers of education. They not only shape education but also use and receive assessment results. Because educational institutions are aware of teachers' unique place in the classroom, they give teachers specific jurisdictions in assessment for various purposes and practices. Hogan (2007) suggests that determining the aim of an assessment can assist in deciding the most suitable type of assessment. Similarly, Taylor and Nolen (2008) stated that teachers can use various assessment methods for classroom goals when they understand what they want their pupils to learn. Moreover, for a complete and sufficient assessment system, teachers should also engage in self-assessments, monitor students' success, learn by observing and measuring, report to parents about their students, or be aware of students' levels (National Research Council, 2001). Therefore, teachers should define and internalize the purposes of the assessment well to ensure that they use assessment effectively. They must then analyze the assessment data for their purposes and make the necessary inferences (Suurtamm et al., 2016). Assessment can serve several different purposes (Barnes et al., 2017; National Research Council, 2001), and there are many tools and methods available for assessment (Black, 2002). Therefore, when selecting or using assessment tools and techniques, teachers should ensure that they are suitable for the intended purpose. According to Black (2002), the purposes of assessment are as follows: (a) supporting learning; (b) certification for ranking, progress, and advancement; and (c) accountability. These purposes are explained in further detail below:

Assessment for supporting learning: According to Black et al. (2003), the primary purpose of assessment for learning (supporting learning) is to facilitate students' learning. Therefore, assessment is one of the basic building blocks of quality education and has an essential place in education. The assessment tools should give

students feedback in a timely and constructive manner on what they have learned and teachers feedback on what they have taught. Thus, both teachers and students conduct self-assessments. Moreover, courses are improved with changes made by feedback from the assessment. In other words, assessment becomes a learning-enhancing tool if it provides the correct data and information to improve the educational process. All assessments (formative, summative, diagnostic assessments, etc.) help teachers recognize students' weaknesses and learn what they need to succeed in the classroom. According to Flórez and Sammons (2013, p. 6), the most frequently cited assessment features for learning in literature are a better use of "questioning, feedback, peer and self-assessment, and the formative use of summative assessment". The use of these features to assess learning has been shown to have positive effects on students, leading to increased success.

Assessment for certification, development, and transition: Students can switch between classes, schools, and teachers throughout their education. While students undergo changes in the environment and person during these transitional stages, conveying the relevant information is crucial. This allows the new teacher to plan his/her teaching effectively and guide each student appropriately (Black, 2002). Altbach et al. (2019) argued that certification is necessary for certain qualifications and achievements that institutions need in order for students to attend university, pursue further education, or find employment after graduation. In other words, assessment data is used to help individuals advance and move to the next learning step. For this reason, there is a need for both institutional and individual certificates (Archer, 2017). From an institutional perspective, certification is confirmation by an accreditation foundation, such as the national education system, that a qualification meets the conditions established by the competent authority. From an individual point of view, a diploma is necessary to certify that individuals have the skills and knowledge required to progress to the next grade level or study.

Assessment for accountability: In many countries, schools are publicly financed, which means that they're responsible for demonstrating that they're facilitating scholar mastering. Governments create responsibility structures that promote effective coaching and studying, boom student success, and inspire lifelong learning to satisfy their accountabilities. Accountability structures typically include placing

gaining knowledge of goals, organizing progress standards, and reporting whether those goals are met. Test outcomes ought to have commonplace standards and standards to advance students' learning to make sure accountability. Therefore, faculties use numerous public well-known assessments, consisting of standardized checks, to measure and record their overall performance, which allows them to be compared to different colleges and establishments on the countrywide or global stage. Additionally, the assessment outcomes ought to be reported honestly and comprehensively to avoid unfair competition among schools and institutions. By the usage of evaluation to increase and improve themselves, schools can fulfill their duty to the public and make a contribution to a more powerful and equitable schooling system (Archer, 2017; Black, 2002; Burger, 2000).

All in all, schools, and institutions train individuals for specific purposes. At the end of the teaching-learning process, an assessment is made to determine whether the goals and behaviors related to the purposes are gained or not. The rate of realization of the objectives at various levels and courses taught in schools emerges from the assessment. A further arrangement is made to enhance and improve the teaching-learning process by using the assessment results. For this reason, it is very critical to determine the aims of the assessment and apply them toward these aims (Bozkurt, 1995).

2.1.2 Standards of Assessment

Teachers must make appropriate decisions for their teaching, using assessment data to understand students' learning levels and promote effective learning (Gullickson, 2003). When assessing or preparing assessment tools, teachers should focus on mathematics, curriculum alignment, and adherence to standards (Suurtamm et al., 2016). According to Gullickson (2002, p.5), standards are defined as "principles mutually agreed upon by individuals involved in the professional practice of evaluation that, when met, enhance the quality and fairness of the evaluation." Yarbrough et al. (2010) stated that these assessment standards help teachers learn about students' learning and to understand whether the desired level of success has been achieved. It provides teachers with a strong framework for thinking about

and organizing assessment. Vocational training organizations began developing teacher competency standards for measurement and assessment in 1987 for the assessments to be more beneficial in terms of improving learning. It was implemented in 1990 when the assessment committee members finished their studies (Gürsoy, 2017). In 1990, the National Council on Measurement in Education (NCME), the National Education Association (NEA), and the American Federation of Teachers (AFT) designed the Standards for Teacher Competence in Educational Assessment of Students. Assessment competencies refer to the knowledge and skills that educators must possess. The deficiency of teachers in evaluating students' progress has highlighted the necessity of establishing standards for assessment competencies. They offered the seven standards considered crucial for evaluating students' academic progress in prospective teacher training and diploma programs. The seven standards are as follows (AFT et al., 1990):

- **Standard 1: Teachers need to possess the competence to select proper assessment techniques, tools, and methods for teaching decisions.** Teachers can use many readily available assessment tools to measure and evaluate their students. Although some of these tools are of good quality, many assessment tools need to be more suitable. Teachers who will measure their students using an existing assessment tool should (a) have knowledge of different assessment methods, (b) know their students well, (c) have knowledge of the curriculum, (d) consider the strengths and weaknesses of each method, (e) should determine the purpose he/she wants to measure and (f) should be able to choose the most appropriate assessment tool for his/her students' level. In addition, teachers need to understand statistical concepts such as reliability, validity, and evaluation error to measure students' progress effectively.

- **Standard 2: Teachers need to possess the competence to generate proper assessment techniques, tools, and methods for teaching decisions.** Teachers may not always be able to find a ready-made assessment tool suitable for their purposes and the level of their students. Hence, they should create a personalized evaluation method instead of relying on pre-made measurement instruments. Developing such tools requires necessary knowledge and skills to determine what they want to measure and design an suitable measurement

tool accordingly. For example, teachers should comprehend the items that will constitute the assessment tool they will develop, know what knowledge and skills they want students to display in the exam, and be aware that the assessment created should appeal to every student in a classroom. Moreover, teachers should be able to gather necessary information from students and use this data to evaluate the effectiveness of the assessment methods they have developed.

- **Standard 3: Teachers need to possess the competence to administer, score, and interpret the assessment results.** Teachers have a critical role in assessment. Selecting appropriate evaluation methods alone is not enough. Applying these tools correctly and efficiently also matters. Teachers should possess remarkable skills to manage, score, and interpret data from various methods while being aware of not increasing students' anxiety levels. They need to consider all the evaluation tools used when judging student performance and be able to interpret results statistically for accurate analysis. In other words, teachers need to be able to read what their results mean statistically. Teachers should also have a knowledge of the rubric types and be able to use them to give points in order to be able to interpret results fairly and adequately.

- **Standard 4: Teachers need to possess the competence to use assessment results for making decisions related to students, instructional planning, curriculum development, and school improvement.** Teachers use assessment results in various situations, such as educational decisions about their students, the school, the aims or results of education, etc. Teachers should be able to transfer all the knowledge and skills they have acquired in assessment to the teaching plans they prepare to realize the educational development of the students. For example, they should recognize this and avoid misinterpretations, if their analyses are invalid according to the applied curriculum. Using the assessment data they collect, teachers should know appropriate uses for improving student, classroom, school, district, state, and national education.

- **Standard 5: Teachers need to possess the competence to create accurate techniques for grading students based on their assessments.** Teachers

consider all assessments made throughout the term to determine an overall grade for students after a training period. Therefore, it is compulsory for educators to fully comprehend grading principles in order to ensure an accurate reflection of student achievements. Teachers are accountable for giving grades based on personal choices and judgments while prioritizing rationality and fairness prior during this process. Providing a system through which students can comprehend their results and accurately report the grading methods used is essential. Moreover, teachers must be able to assess their processes impartially and alter them when needed, thereby improving student performance analysis validity.

- **Standard 6: Teachers need to possess the competence to explain assessment findings to students, colleagues, and parents.** Teachers should keep students and parents informed and continuously report on the assessment results. Moreover, these results should be discussed with other educators. In this way, the students, teachers, and other educators constantly communicate with the teacher. So, the assessment results are understood in detail while misunderstandings are prevented. Thus, they can defend their evaluation analysis and interpretation when they need to defend themselves. Teachers should be competent in explaining how assessment results can be interpreted according to socio-economic, cultural, linguistic, and other factors. They send printed assessment results to the classroom, school, state, etc., and they should be submitted where necessary.

- **Standard 7: Teachers need to possess the competence to identify assessment information that is unethical, illegal, or otherwise inappropriate.** Teachers should ensure the rules of ethics and the rights of everyone involved, namely justice while collecting information, interpreting the information they collect, using it, and transmitting it, starting from the first planning of all their evaluation processes. Teachers should be familiar with laws and court decisions that affect assessment and assessment practices in the classroom, school, and state. Teachers should know that the assessment results should not be abused or overdone; that will embarrass students and violate their rights.

The standards created by AFT, NCME, and NEA in 1990 have significantly shaped teacher education. These standards have influenced the design of course content, as discussed by O'Sullivan and Johnson (1993). They serve as a valuable tool for teachers' self-assessment and support educators in developing a deeper understanding of student assessment in educational measurement (Brookhart, 2011). Furthermore, these standards have helped define the competencies required for teachers and administrators in various educational fields (Mertler & Campbell, 2005; Plake et al., 1993). While the following researchers attempted to create standards like those established by AFT, NCME, and NEA in 1990, none achieved the same widespread adoption level (McMillan, 2001; Stiggins & Chappuis, 2008). In 2011, Brookhart expanded to the 1990 standards and developed an updated version of 11 items. However, the assessment literacy inventory developed by Mertler and Campbell in 2005 based on the 1990 standards is used as a data collection tool for this study.

2.1.3 The Process of Assessment

Nitko and Brookhart (2007) highlight the importance of teachers considering multiple factors, including curriculum content, instructional strategies, and assessment methods, when planning their teaching and evaluating students' learning. They emphasize the need for comprehensive data collection and rigorous assessments to make well-informed and effective decisions about student performance. Teachers who possess a strong understanding of assessment can seamlessly integrate it into instruction, employing appropriate teaching methods (McMillan, 2003). Similarly, Thomas et al. (2004) stress the significance of teachers gathering diverse forms of assessment data, tracking student progress, monitoring learning outcomes, and collaborating with colleagues, parents, or students to stay informed about educational progress. Assessment encompasses the range of information-gathering practices used by teachers to determine suitable instructional methods and support student learning. Consequently, assessment findings significantly influence decision-making processes (Nitko & Brookhart, 2007; Thomas et al., 2004). Furthermore, the collection and analysis of information are integral to enhancing teaching and learning outcomes through the assessment process. Therefore, accurate and effective execution of each step of the assessment procedure is crucial (Simeonsson & Bailey,

1988). Nitko and Brookhart (2007, p.14) propose a comprehensive framework for the assessment process, comprising five key responsibilities: (1) generating or choosing the assessment; (2) administering the assessment; (3) scoring the assessment results; (4) interpreting and utilizing the assessment results; and (5) communicating assessment results.

2.1.3.1 Generating or Choosing the Assessment

The first step in the assessment process involves creating assessments that effectively improve students' learning. Teachers must determine the purpose of assessment when selecting or designing assessments that will enable students to learn better. In other words, it is important to carefully choose assessment methods, items, and tools that align with the intended assessment goals (Haladyna & Downing, 2006) since the data provided by the evaluation created for a specific purpose is more useful and effective (Newton, 2007). Moreover, assessments must be valid and reliable (CSAI, 2018). Without validity and reliability, teachers cannot make informed decisions or effectively plan their instruction (Taylor & Nolen, 2008). To be considered valid, an assessment tool must accurately measure the intended property while minimizing the impact of other variables (Darr, 2005). The more accurate an assessment tool is in assessing the particular trait it aims to measure, the more trustworthy it is (Atılgan et al., 2007; Haladyna & Downing, 2006). Validity of an assessment means that the deductions and conclusions made by teachers founded on the results of the assessment are based on solid foundations (Nuttall, 1987). Consequently, teachers should prioritize alignment with learning goals and objectives for classroom assessments, ensuring that the results positively contribute to students' development (Nitko & Brookhart, 2007). Furthermore, assessments should consistently and accurately evaluate students' abilities across different testing situations and teachers to ensure validity for all students (Russell & Airasian, 2012).

On the other hand, reliability refers to the consistency of the outcomes that teachers get from an evaluation (Darr, 2005). The reliability of assessment methods is associated with how consistent the results are when using the same assessment tool under similar conditions (Atılgan et al., 2007). In other words, reliability refers to

the consistency of students' scores on the same or similar tasks, and it is crucial to examine the consistency of scores across multiple assessments to assess the reliability (Nitko & Brookhart, 2007). Darr (2005) stated that although it is difficult to achieve complete reliability, it is vital to have a maximum reliable assessment tool to make accurate inferences. Because reliable assessment results help teachers accurately determine students' level of achievement and make educational decisions while also providing feedback to students, parents, and educators. On the other hand, unreliable assessment results can lead teachers to incorrect decision-making in their teaching (Russell & Airasian, 2012). To ensure accurate judgments, teachers should employ as many reliable assessment methods as possible (Taylor & Nolen, 2008).

2.1.3.2 Administering the Assessment

Once teachers have chosen or created the assessment to apply to their students, it should be implemented so that all students in their class can fully demonstrate their knowledge and skills. In other words, teachers should create a physical and mental atmosphere for students that allows students to do their best (Russell & Airasian, 2012). For example, assessments should be meaningful and motivating for students to encourage them to perform their best (Sanders & Horn, 1995). Therefore, teachers are responsible for ensuring that assessment procedures are applied equally to all students (Elturki, 2020). Besides teachers, students also need a basic understanding of assessment management to perform well in assessments (Nitko & Brookhart, 2007). According to Nitko and Brookhart (2007), students need to be aware of the assessment date and time, the context and requirements of the assessment, the skills and content that will be tested, the emphasis of the assessment, the standard or criterion of performance that will be used to score, and the consequences of the results in decision-making.

Moreover, teachers should avoid creating an atmosphere that threatens their students and ensure that the environment in which the assessment will be conducted is peaceful and respectful. In addition, the environmental conditions in which the assessment will be applied should be fair for each student and not distracting (Russell & Airasian, 2012). Saad et al. (1999) noted that teachers should not create test anxiety in students

because this causes to be an obstacle in showing students' performance and real success. The various responsibilities that teachers must fulfill before and during the exam are: (a) being present in the classroom before the exam starts, (b) checking the classroom environment and table setting, (c) giving students enough time to read the instructions and questions, (d) taking the necessary precautions to prevent cheating and (e) informing students about the exam period (Elturki, 2020).

2.1.3.3 Scoring the Assessment Results

After the assessment tool is applied, the assessment results are scored to have data to be interpreted (Saad et al., 1999). Measurement is crucial to test scoring, as it assigns numerical values to represent a student's performance. Regarding accomplishment tests, how well a student does on the test items is translated into a score used to determine student performance decisions (Russell & Airasian, 2012). The factors that teachers should pay attention to while scoring the evaluation results are as follows (Elturki, 2020; Nitko & Brookhart, 2007): (a) to make an accurate and objective assessment, tools such as rubrics, assessment scales, or keys should be used, (b) the scoring process for students should be free from any factors that could lead to unfair outcomes, ensuring that all students are evaluated relatively, (c) scoring anonymously, one approach to achieve greater objectivity is to either grade together as a team or assign specific sections for each teacher to grade, (d) providing feedback on why students' wrong answers are wrong to improve their learning, to explain the rubric used and the logic behind correct answers, and allow students to have access to their assessments for review. A teacher who wants to further students' learning should regularly evaluate their students and score the assessment results correctly (Taylor & Nolen, 2008). For this reason, it is essential for the assessment process to properly progress the procedure required for scoring while teachers are scoring the assessment results (Atılgan et al., 2007). Without accurate scoring, teachers cannot make reasonable and appropriate decisions for teaching, as an incorrectly scored assessment cannot be appropriately interpreted (Russell & Airasian, 2012).

2.1.3.4 Interpreting and Using the Assessment Results

Teachers create a teaching process according to the learning goals they want their students to acquire. Teachers should constantly monitor the student's learning process and make some inferences to understand whether these learning objectives are earned or not (Nitko & Brookhart, 2007). They need to understand the data they collect from students, that is the assessment results, for teachers to make reasonable inferences. Therefore, one of education's most critical components is to interpret and use the evaluation results (Anderson, 2003). According to Nitko and Brookhart (2007), what is expected from a good teacher is to analyze and interpret the assessment results they collect from students in a meaningful way, both individually and on a class basis. To analyze the evaluation results well, they need to make comparisons by taking into account different data and making a holistic interpretation (Downing, 2003). Therefore, it is necessary to use different assessment methods that will enable them to make more than one comparison (Anderson, 2003). According to Taylor and Nolen (2008), there are two types of interpretation and comparison of the data obtained as a result of the evaluations: "Criterion Referenced Interpretation" and "Norm-Referenced Interpretation." In criterion-based interpretation, a criterion is determined for achievement, and the student is considered successful if he or she can meet this criterion. For example, if the success criterion for an exam is determined as getting at least 50 points, students who score 50 and above are considered successful. A well-done criterion-based assessment is beneficial to the teacher in identifying students' strengths and weaknesses for improvement (Downing, 2003).

On the other hand, it is not clear in advance in which situation students will be considered successful in norm-referenced interpretation because students' success depends on the group in it (Mehrens & Ebel, 1979). In other words, the students' achievements are compared with their classmates, and as a result of this comparison, a comment is made on whether they are successful or not. For instance, interpretations made according to the mean or curve fall under norm-references interpretation. With norm-referenced interpretation, teachers can rank their students' achievements and have an idea of the performance of all students at the grade level (Shepard, 1979). Teachers should choose the type of interpretation according to their purposes.

However, whatever type of assessment is used, teachers must use the information they have obtained due to interpretation to advance students' learning (Russell & Airasian, 2012).

2.1.3.5 Communicating the Assessment Results

Teachers must collaborate with students, parents, and other educators who impact students' learning before changing their teaching (Anderson, 2003). They should frequently communicate with them and provide regular student progress reports (Nitko & Brookhart, 2007). In addition to being in constant communication, it is also crucial that communication be open and constructive in order to improve students' learning. In parallel with this, Deneen and Deneen (2008) also emphasized that only the teacher's effort is not enough to develop students' learning, and students, parents, and other educators should also fulfill their responsibilities.

Teachers must ensure regular feedback on students' tasks and performances because providing feedback helps enhance students' understanding, who often seek knowledge about their performance (Russell & Airasian, 2012). Teachers can provide valuable feedback on student strengths and weaknesses for improvement by providing written comments on student assignments. Alternatively, they can negotiate with students regularly to share their thoughts on their performance during the lesson Stiggins (1994). In addition, they can review assessment results with students and explain how the assessments were scored. This way, the learning process is positively affected because students can understand where and how they made mistakes (Brookhart, 1999). On the other hand, teachers need to communicate with parents and other educators about how students are progressing toward learning goals. This is essential as it ensures parents' and other educators' assistance to teachers (Blaich & Wise, 2011). Moreover, Taylor and Nolen (2008) emphasized that communication should not be one-way, like only from teacher to parent or educator. Parents and other educators should also know that they must convey information to teachers.

2.2 Assessment Literacy

Literacy infers an individual's competence in a particular field (Gee, 2005). According to the definition made by Organisation for Economic Co-operation and Development (2006), an individual's capacity to advance him/her knowledge and abilities to actively engage in society, grasp and successfully use the textual resources that are readily available, and develop well-informed judgments is referred to as literacy. Many definitions of literacy, such as technology, computer, environmental, science, math, or art, have been developed in literature (Wagner, 2009). The concept of literacy has become an important concept in the field of measurement and assessment in recent years (Popham, 2004). Assessment literacy, accepted as a critical component of pedagogical content knowledge (Shulman, 1986), is an essential professional knowledge and skill that teachers should acquire (Volante & Fazio, 2007). Assessing students' performance is a crucial task for classroom teachers, and the quality of teaching is directly linked to the quality of assessments used, according to Stiggins (1999). Therefore, the data from classroom assessments must be valid and reliable to be significant and proper. To achieve this, teachers need to be knowledgeable in assessment, which is referred to as being assessment literate (Mertler & Campbell, 2005).

There are many definitions in the literature on assessment literacy, and it was first introduced and defined by R. J. Stiggins (1991). Stiggins (1991) argues that assessment-literate teachers must be able to choose what to measure; create appropriate measurement tools to assess student achievement; provide examples for demonstrating student success; and address any issues or challenges that may arise to prevent them from occurring again. Similarly, DeLuca and Klinger (2010) declared that educators, who have a skill in assessment literacy, know what, why, and how to evaluate learning. They are aware of the problems regarding assessment and wise how to prevent them from occurring. Moreover, they also understand the possible negative consequences of an inadequate and incorrect assessment. Assessment literacy also includes creating trustworthy assessments and then managing and scoring those assessments to foster the right teaching decisions based on state or provincial education standards (Xu & Brown, 2016). Popham (2004) stressed the significance

of assessment literacy for teachers and referred to the lack of knowledge in this area as "professional suicide." He further noted that an assessment-literate teacher should understand basic assessment concepts and procedures and should be able to understand and interpret test results accurately.

Gottheiner and Siegel (2012) stated that teachers with assessment literacy could interpret the data obtained from the assessment results, communicate with students about their performance, and use the information obtained from the analysis for long-term and short-term learning purposes. McMillan (2001) emphasized that assessment-literate teachers successfully combine the teaching process with assessment and therefore use the most suitable teaching techniques for their students. Teachers who can integrate the teaching process with assessment have essential contributions to teaching, such as effective learning of students, more quality lessons, and higher education standards (DeLuca & Klinger, 2010). Mertler (2004) stated that an assessment literate teacher has the following characteristics:

- Understanding which assessment method teachers should use to measure students' learning and achievement reliably.
- Using assessment to increase students' motivation and learning by evaluating, keeping reports, and constant communication with them.
- Effectively analyze and interpret assessment results using reports, exams, achievement tests, portfolios, or conferences.

Newfields (2006) attributed the importance of assessment literacy to three reasons. First, assessment is a standard part of all education systems. At the same time, teachers devote much of their time to assessment, and schools allocate significant budgets for evaluation. To not waste all these efforts, teachers must have assessment literacy for high-quality assessment. Second, teachers with assessment literacy grasp basic statistical concepts. Thus, research does not stray away from the reality of science; fundamental knowledge emerges. Therefore, it provides an understanding of the literature on education. This becomes a tool for understanding how the teaching programs work. Finally, being assessment literate allows teachers to communicate results about the general state of the classroom to others. Thus, the teacher shares

their data with other colleagues and results that encourage learning. It also enables teachers to self-criticize by comparing their performance with other teachers. For this reason, the teacher's assessment literacy significantly contributes to education.

Although assessment literacy is crucial for quality education, most teachers do not comprehend the fundamental concepts and procedures of assessment and thus are not assessment literate (Brookhart, 2001; DeLuca & Klinger, 2010; Mertler & Campbell, 2005; Plake et al., 1993; Popham, 2004; Stiggins , 1995; Volante & Fazio, 2007). Chappuis and Stiggins (2002), Clark et al. (2022), and Ye (2023) stated that the assessment literacy levels of teachers are inadequate, which results in the improper evaluation of students and prevent them from achieving their full potential. According to Lai and Waltman (2008), the assessment tools that teachers who lack assessment literacy will use to evaluate student achievement cannot provide construct validity and reliability. In this situation, the teachers will fall victim to systematic rather than random error, endangering the educational system. According to the literature, many factors affect teachers' low assessment literacy. These factors are; that teachers do not receive sufficient education on effective assessment during their undergraduate education (DeLuca & Klinger, 2010; Mertler & Campbell, 2005; Schafer, 1993; Stiggins , 1995; Volante & Fazio, 2007), the lack of awareness about the need to be literate in assessment (National Research Council, 2001), the pressure of education policy on the use of assessment based on summative and accountability (Hill, 1999), school administrators' assessment literacy levels are insufficient (Popham, 2004; Stiggins , 1995), teachers' lack of content knowledge (Ogan-Bekiroglu, 2009), the fact that the faculty members whom the pre-service teachers take assessment courses do not have full knowledge of the field (Heritage, 2007), wrong or inappropriate perspectives and attitudes towards assessment (Beck, 1980; Stiggins , 1995). These factors affecting assessment literacy can be described as barriers to teachers' being assessment literate. However, since assessment literacy is crucial to improving students' learning by improving the quality of teaching and learning (Volante & Fazio, 2007), and assessment literacy should be emphasized more (DeLuca & Klinger, 2010).

2.2.1 Assessment Literacy and Teacher Education Programs

Xu and Brown (2016) noted that assessment literacy is increasingly vital to teacher professionalism. This increased attention to assessment literacy is due to recognizing assessment's significant impact on student learning (Popham, 2018) and realizing that teachers play a crucial role in assessment (Plake, 1993). Stiggins (1995) further emphasized the importance of teachers having sufficient assessment literacy to support students' higher academic achievement. Given that teacher candidates are inexperienced teachers who just have finished their teacher education programs and are about to begin their instruction careers, providing high-quality assessment education during the pre-service period is essential to building a solid foundation for future professional development (DeLuca & Klinger, 2010). Although the literature on the importance of assessment literacy in education agrees, many teachers and teacher candidates lack proficiency in assessment, and their assessment literacy levels are low (Schafer, 1993). Many teachers fail to meet the assessment competencies outlined by the AFT, NCME, and NEA (1990) because they cannot effectively integrate assessment into their teaching. Plake and Impara (1993) linked the insufficient pre-service training of teachers to their lack of comprehension of proper evaluation instruments and processes. According to a study by Beziat and Coleman (2015), the assessment literacy levels of pre-service teachers were not improved by the courses they took as part of their education. In research conducted by Mertler (1999) revealed that pre-service teachers who completed assessment courses still lacked the ability to evaluate students effectively by the end of the course. Unfortunately, many teachers were not required to learn concepts about educational assessment when they finished their undergraduate education (Popham, 2018). Considering the teacher education program that teachers have taken, it is understood why they have such an extensive knowledge gap in assessment issues.

Despite there is a great need for teacher preparation programs, there are still very few situations where formal assessment training is compulsory (DeLuca & Klinger, 2010). For instance, based on research conducted by DeLuca and McEwen (2007), only three of the ten Bachelor of Education programs in Ontario offered a course in assessment. Similarly, Volante and Fazio (2007) stated that only two states in the U.S.

and Canadian provinces, out of 50 and 10, allocate substantial funding to enhance school assessment procedures. On the other hand, (Wise et al., 1991) discovered that teachers' pre-service training included less than one period of learning session on assessment-related topics, which resulted in a lack of knowledge and proficiency in assessment and associated methods. Although assessment is now given greater importance in schools and colleges in the U.S., many of these institutions still do not make it mandatory for pre-service teachers to take undergraduate courses that focus on classroom assessment (Mertler & Campbell, 2005; O'Sullivan & Johnson, 1993). Similarly, in studies conducted in Turkey, it has been revealed that teacher candidates' taking measurement and assessment courses, or in-service training does not change their assessment competencies (Çelikkaya et al., 2010; Eğri, 2006; Kuran & Kanatlı, 2009). The insufficient education on assessment during teacher training and the absence of research on effective teaching strategies in this field have been identified as key elements causing the low assessment literacy levels (DeLuca et al., 2013; Graham, 2005; Popham, 2004). The lack of opportunity for teachers to learn about classroom assessment has resulted in few of them being adequately prepared to handle the associated challenges (Stiggins , 2004). Volante and Fazio (2007) argue that there is still a limited amount of research dedicated to exploring the assessment literacy of teachers, and this research should start in faculties of education, as these institutions offer prospective teachers their initial exposure to assessment. Considering the scope of measurement and assessment in Turkey, in the document "Education Faculty Teacher Training Undergraduate Programs" published by the Yüksek Öğretim Kurulu (YÖK) (2007), a 3-credit theoretical assessment and evaluation course is included in the curriculum of all teacher education programs in education faculties. The content of the measurement and assessment course included in teacher education programs is as follows (Yüksek Öğretim Kurulu (YÖK), 2007):

- The place and importance of measurement and evaluation in education,
- Basic concepts of measurement and evaluation,
- Qualities that need to be found in assessment tools (reliability, validity, usefulness),
- Assessment tools and features used in education,

- Assessment tools based on traditional approaches (written exams, short-answer exams, true-false type tests, multiple-choice tests, matched tests, oral exams, assignments),
- Assessment tools for getting to know the student in multiple ways (observation, interview, performance evaluation, student product file, research papers, research projects, peer assessment, self-assessment, attitude scales),
- Basic statistical operations on assessment results,
- Evaluating learning outcomes,
- Grading,
- Developing an assessment tool.

Despite a detailed offered regarding the content of the measurement and evaluation course published by YÖK (2007), it has been observed that the studies do not give enough importance to the measurement and evaluation issues in undergraduate education (Baykul, 2000) and that the pre-service teachers cannot achieve the determined objectives of the course significantly (Çelik & Önal, 2005). For this reason, it was stated that teachers who graduated from teacher education programs were insufficient in assessment (Çakan, 2004).

All in all, education faculties are most crucial in preparing teachers to conduct measurement and assessment. As Volante and Fazio (2007) stated, education institutions that prepare future teachers should successfully acquire knowledge, skills, and attitudes linked to assessment, like assessment principles, assessment tools and techniques, data analysis and interpretation, fair and objective evaluation, feedback, etc. However, earlier research has indicated that educators don't feel fully equipped to properly evaluate their students' progress (Ogan-Bekiroglu & Suzuk, 2014). Although it is still insufficient, studies have been carried out to shed light on academicians to train more qualified teachers in the evaluation of the education given in teacher education programs. Studies on the assessment literacy levels of pre-service and in-service middle school mathematics teachers are explained in more detail in the section below.

2.2.2 Studies on Assessment Literacy

A teacher must be proficient in assessment to rule the classroom and judge students' learning performance (Beziat & Coleman, 2015). Even though assessment literacy is crucial for quality education, Popham (2018) stated that most teachers do not comprehend the fundamental concepts and procedures of assessment and, thus, are not assessment literate. According to Campbell et al. (2002), one of the biggest reasons for the low assessment literacy of teachers is that they cannot take adequate and effective courses about assessment in undergraduate teacher education. Mertler (2003) stated that what teachers learn in these courses does not match what they need to know in classroom practice. Therefore, he argued that educators should be aware of the importance of assessment courses and develop more efficient courses in the teacher education program. Education faculties should know the necessary teacher professional development opportunities, so prospective teachers who complete teacher education programs can graduate more competently in assessment (Mertler & Campbell, 2005). For this, appropriately designed measurement tools are needed to determine the level of knowledge about assessment literacy (Bütüner et al., 2010). The scales developed to measure teachers' assessment literacy and the qualitative and quantitative studies conducted in this field are mentioned below in more detail.

2.2.2.1 Developing Assessment Literacy Inventory

Some research was done in the literature, and scales were developed to measure assessment literacy. Plake et al. (1993) developed Teacher Assessment Literacy Questionnaire (TALQ) to measure teachers' assessment literacy levels. While preparing the questionnaire, the researchers considered seven standards specified by AFT, NCME, and NEA in 1990 as a guiding framework. They developed a questionnaire comprising 35 content-based items containing five items for each standard. The researchers applied this questionnaire they created to 555 in-service teachers. They calculated the internal consistency coefficient (K-20) as 0 .54. On average, teachers answered 23 of 35 items correctly. Researchers interpreted the participants' correct answers to 23 of 35 questions as teachers' literacy levels are low and teachers should be more qualified to accurately measure students' learning.

Moreover, the researchers compared the assessment literacy levels of teachers who took and did not take assessment courses. In addition, they stated that teachers who took assessment courses had higher assessment literacy levels.

Campbell et al. (2002) simply gave a new name to the Assessment Literacy Inventory (ALI) without touching the items in the questionnaire developed by Plake et al. in 1993. In other words, the researchers used the same scale created by Plake et al. (1993) in their study by simply changing its name. Unlike the study by Plake et al. (1993), they applied the inventory to prospective teachers to measure their assessment literacy level. They administered the inventory to 220 undergraduate students enrolled in the teacher education program and completed the assessment courses. The reliability coefficient of this scale (0.74) was higher than Plake and Impara's (1993) study on in-service teachers. The average number of correct answers in 35 items of pre-service teachers was 21. Therefore, they stated that pre-service teachers (M=23) answered two questions more correctly than pre-service teachers (M=21) compared to the study by Plake et al. (1993). The researchers also noted that both groups scored the lowest in the sixth competency standard, which involves communicating assessment results to various parties such as students, parents, and other educators.

Mertler (2003) also conducted a study about assessment literacy by using the Classroom Assessment Literacy Inventory (CALI), which was the minor revised version of the questionnaire created by Plake et al. in 1993. Like Plake and his colleagues (1993) and Campbell et al. (2002), Mertler (2003) also desired to investigate the assessment literacy levels of participants. However, unlike the other two studies, Mertler (2003) aimed to measure the assessment literacy of both pre-service and in-service teachers and compare the results of these two groups. Participants of the study were 197 in-service and 67 pre-service science and social studies teachers. The findings revealed that pre-service teachers correctly answered approximately 19 out of 35 questions. When their scores were examined based on competency areas, the highest score was achieved in the first competency area, which is choosing suitable assessment methods. The lowest score was recorded in the fifth competency area, which pertains to developing a valid grading process. In contrast, in-service teachers answered less than 22 questions accurately according

to study findings. Furthermore, the highest score of in-service teachers was in the third competency area, which is administering, scoring, and interpreting assessment results. The lowest score was in the fifth competency area, like teacher candidates. The study also found significant differences between the two groups in five of the seven standards, as analyzed by an independent t-test.

The studies conducted by Campbell et al. (2002) and Mertler (2003) were essentially replications of the Plake (1993) study, as they utilized the same assessment literacy instrument developed by Plake. However, both studies discovered that the instrument had inadequate psychometric properties when administered to both pre-service and in-service teachers. Additionally, the instrument was lengthy, difficult to comprehend, and presented questions in a disjointed manner. Consequently, Campbell and Mertler (2005) proposed a complete revision or redevelopment of the assessment literacy instrument. To accurately measure teachers' assessment literacy, the researchers embarked on the development of a new assessment literacy tool. They conducted a two-stage pilot study, involving 152 teacher candidates in the Fall of 2003 and 249 teacher candidates in the Spring of 2004, to test the instrument. The Assessment Literacy Instrument (ALI) exhibited robust psychometric properties, indicating that it is a valid measure of educators' assessment literacy.

In 2016, DeLuca et al. argued that the existing assessment literacy scales in literature do not meet today's standards and aimed to develop a new tool, considering the important changes in the assessment field. While developing their scales, researchers used Classroom Assessment Standards (Joint Committee for Standards in Educational Evaluation, 2003) as assessment standards. The participants of the study are 150 teacher candidates and 254 teachers. The researchers named their scale "The Approaches to Classroom Assessment Inventory" ACAI for short. The researchers gathered their data through a survey and an email invitation shared on numerous social media and media networks. The researchers created a three-part instrument to assess teachers' proficient knowledge and assessment practices: "(1) approaches to classroom assessment, (2) confidence in classroom assessment, and (3) priorities and preferences for professional assessment learning." According to the study's findings, the scale developed by the researchers passed the validity and reliability test. Additionally, they claimed that the ACAI offers a brand-new and essential tool

to enhance assessment literacy research and professional growth.

2.2.2.2 Pre-service and In-service Teacher's Assessment Literacy

The study conducted by Volante and Fazio (2007) aimed to determine pre-service teachers' perceptions regarding assessment literacy, the purposes for which they used assessment, the assessment tools they employed, and suggestions for improving the assessment process. An annual survey was applied to the study group of 69 pre-service teachers conducting a teaching experience course in their education programs. The analysis revealed that the pre-service teachers had low self-efficacy perceptions in assessment and preferred summative assessment over formative assessment. They expressed a more critical requirement for additional pedagogy in authentic assessment methods. Moreover, participants strongly advocated the improvement of a particular lesson(s) concentrated on classroom assessment to increase their understanding of assessment. According to the researchers, the study's results indicated that the teacher education program did not adequately cover different classroom assessment techniques. As a result, they recommend that the program should give greater awareness of a broader range of assessment methods mentioned in the course main line. Lastly, the main recommendation is to choose mentor teachers who can demonstrate proper classroom assessment abilities to enhance teacher candidates' assessment literacy in teaching practice environments.

DeLuca and Bellara (2013) aimed to examine (1) the consistency between policies for teacher education accreditation, (2) professional standards for teacher assessment practice, and (3) the content of assessment courses of the teacher education program. Two policy papers, two professional standards papers, and the ten Florida-based curriculums were used for curricular alignment. They used a medium complex alignment approach to identify areas of agreement and disagreement among policies, standards, and curriculum elements related to teacher assessment competency development. The study found both areas of agreement and disagreement among the policies, standards, and curricula analyzed. The researchers suggested greater alignment between pre-service assessment education and professional standards and recognizing assessment literacy as a fundamental teaching skill.

McGee and Colby (2014) conducted a study aimed to examine how an assessment course affected the assessment literacy of teacher candidates; to identify the areas where the teacher candidates showed the most improvement in their assessment literacy; and to determine whether there were differences in assessment literacy levels among candidates in different licensure areas. The researchers picked five distinct license levels and 19 different teacher education majors among the participants. They used pre-test and post-test designs with 190 teacher candidates enrolled in an assessment course. The "Assessment Literacy Inventory," developed by Mertler and Campbell (2005), was used as a data collection tool in the research. When the scores of the pre-service teachers from the applied inventory were examined, it was revealed that the highest score was on ethical evaluation, scoring, and choosing evaluation methods. The lowest score was on creating assessments and communicating the assessment results. According to the study's findings, pre-service teachers' achievement in subjects with low scores did not improve significantly after taking the assessment course. While there was a difference between the pre-test and post-test scores of the groups, the effect sizes were small. Furthermore, it is concluded that the assessment course was insufficient to improve scores in the areas where candidates scored poorly. Researchers underlined that these two areas should be focused more on within the scope of the assessment course.

Using quantitative and qualitative research methods, Ogan-Bekiroglu and Suzuk (2014) conducted a study to evaluate pre-service physics teachers' assessment literacy and practical application. The quantitative part of the study aimed to measure the assessment literacy levels of the participants who took the assessment course, while the qualitative part aimed to confirm the quantitative findings and explore the participants' evaluation practices in the participants' natural classroom environments. 28 teacher candidates participated in the study. The researchers created a forty-four-item scale based on McMillan (2001) scale and used it as a data collection tool. An evaluation form which contains 11 open-ended questions, and an interview form was also used to assess the homework process of participants. The quantitative results of the study showed that pre-service physics teachers had a relatively high level of assessment literacy, which was confirmed by the qualitative analysis results. However, the researchers discovered a discrepancy between theoretical

assessment literacy and its practice application. The study proposes that teacher education programs should prioritize the comprehension of assessment theories and methods, emphasize the significance of validity and reliability in assessments, and allow students to gain experience in traditional and performance-based assessment techniques. Researchers also underlined the need for pre-service teachers to think more about assessment practices and to provide experiences to improve their skills in this subject.

Beziat and Coleman (2015) aimed to determine preschool, elementary, and secondary education teacher candidates' assessment literacy levels, observe their assessment practices, and propose a program aligned with the deficiencies encountered. The study was conducted with 39 prospective teachers enrolled in various departments at a university in southern England, but only 26 students finished both implementations. The researchers used the Teacher Assessment Literacy Questionnaire (TALQ) developed by Plake and Impara (1992), which consists of 35 multiple-choice questions to measure the pre-service teachers' assessment knowledge. Researchers observed pre-service teachers' understandings of assessment literacy from the first to the last lesson of the assessment course. They stated that pre-service teachers could not comprehend assessment literacy during this process. In other words, the study's findings showed that taking an assessment course had no discernible impact on students' assessment literacy. However, it was revealed that prospective teachers admitted to majors and professional programs in secondary education had a better understanding of assessment procedures at the end of the study. The researchers remarked that it is necessary to investigate whether the content of the assessment courses complies with the assessment standards. Accordingly, necessary adjustments should be made to the course content. They also advise that real assessment experiences should be included in the lessons.

DeLuca and Klinger (2010) conducted a study to explore perceived levels of prospective teachers' confidence in assessment practice, theory, and philosophy. They implemented an assessment literacy questionnaire for 288 pre-service teachers with different branches and grades. The researchers asked the teacher candidates to share their thoughts on the assessment subjects that should be covered in the assessment course in undergraduate education. Also, they measured the confidence levels of

pre-service teachers using a questionnaire. According to the findings of the study, in order for pre-service teachers to acquire confidence in assessment, they need to gain practical experience in courses such as internships, in addition to the assessment course where theoretical knowledge is learned. The study's results demonstrated that the most noteworthy differences in pre-service teachers' confidence levels were seen in their practical and theoretical assessment knowledge rather than philosophy. Hence, the researchers articulated that the assessment course significantly influences pre-service teachers' confidence in assessment practice and theory while having a minor impact on their assessment philosophy. They also suggested that the assessment course covers evaluating and reporting success, ensuring validity and reliability, and developing measurement tools.

In her research study conducted by Gu (2014), an investigation was carried out to examine the approach of English as a Foreign Language (EFL) teachers towards assessments in a Chinese secondary school. By utilizing qualitative research techniques ranging from document analysis to classroom observation and conducting interviews with instructors themselves - data was collected on aspects influencing their approach to evaluations. The outcomes revealed that an array of contextual elements such as institutional policies or even individual preferences impacted how EFL instructors carried out assessments within their classrooms. Emphasizing this finding's significance; Gu advised building efficient evaluation procedures by comprehending these variables while also emphasizing training programs geared towards enhancing educators' "assessment literacy."

Tünkler (2019) examined assessment literacy levels and self-efficacy perceptions of 90 social studies teacher candidates by considering various variables. To collect data for the study, the researcher employed two tools: the "Assessment Literacy Inventory" and "Assessment and Evaluation General Competence Perception Scale." The results indicated that pre-service teachers displayed a deficiency in assessment literacy levels but possessed moderate self-efficacy perceptions. Despite having confidence in executing assessments, participants lacked competence required for effective implementation as pointed out by the researcher. However, results revealed no considerable gap between genders or academic achievements of participants concerning their assessment self-efficacy perceptions. The researcher noted a

remarkable lack of relationship between the participants' assessment literacy levels and their assessment self-efficacy perceptions. She stressed that this indicates the need for additional research to better understand the complex factors contributing to developing effective assessment practices among prospective teachers.

Using a framework that encompassed assessment principles, techniques, and aims, Siegel and Wissehr (2011) examined the assessment literacy of 11 pre-service secondary science teachers. They studied the participants' grasp of assessment tools and justifications for utilizing assessment by gathering information from their journals, teaching philosophies, and inquiry-based science lessons. Additionally, they examined how future teachers included assessment in their inquiry-based science lessons. According to the study, the teacher candidates demonstrated a firm grasp of the principles of assessment for learning. They recognized the value of employing various assessment tools, naming 19 distinct ways to evaluate student understanding. They also utilized assessments for learning purposes in their inquiry-based science lessons. However, the researchers noted a marked contrast in the pre-service teachers' perception of assessment. This contradiction was because the pre-service teachers' teaching philosophies and the theoretical knowledge they expressed in their diaries did not match their practicality when evaluating students.

In his thesis, Gül (2011) conducted a study to assess the assessment literacy levels and attitudes of pre-service primary school teachers specializing in science (n=54), social studies (n=55), Turkish (n=43), and mathematics (n=28) towards assessment. The "Teacher Assessment Literacy Questionnaire" developed by Plake and Impara (1993) was used to assess the assessment literacy of prospective teachers. An attitude scale developed by the researcher following interviews with pre-service teachers was used to determine their attitudes toward assessment. The findings revealed that the assessment literacy levels of the teacher candidates were relatively low. The science teacher candidates had an average of 17.57 correct answers out of 34 questions in the assessment literacy inventory; the social studies teacher candidates had an average of 14.51 correct answers, the mathematics teacher candidates had an average of 20.21 correct answers, and the Turkish teacher candidates had an average of 17.83 correct answers. The assessment literacy standard where the teacher candidates struggled the most was communicating assessment results to students, parents, and other educators.

Additionally, significant differences were observed among the prospective teachers regarding standards.

Yamtim and Wongwanich (2014) conducted a study to explore primary school teachers' assessment literacy scores and to provide strategies for improving this literacy level. The researchers used two data collection tools: "a classroom literacy evaluation questionnaire" and "a focus group discussion methodology." The study involved 8 teachers in a focus group discussion and 19 primary school teachers who finished the "Classroom Assessment Literacy Questionnaire" (Mertler, 2003). The findings showed that the average number of correct answers in the inventory of the participants was approximately 17.11. Thus, many teachers had insufficient proficiency levels in assessment literacy. When the assessment literacy scores are examined in more detail, the best standard of the participants is the first standard, with an average of 3.11 (selecting the proper assessment method). The standard in which the participants received the lowest score was the fifth (valid grading), with an average of 1.79. The researchers suggested that teachers must engage in collaborative learning and have well-informed mentors or trainers who guide teaching practice to enhance teachers' assessment literacy.

Deneen and Brown (2016) desired to determine (a) the perception of assessment concepts among teacher candidates and experienced teachers in a master program, (b) the shifts in these perceptions resulting from the assessment course, and (c) the potential correlation between participants' evaluation visions and assessment literacy. Data on the academic performance of 30 of the 32 students enrolled in the assessment course were examined, and pre-and post-course interviews were conducted with 6 of these 30. The study's findings demonstrated that the participants' perceptions of assessment were favorable and unfavorable and that these perceptions were irrespective of their academic achievement. According to the researchers, while the participants' assessment literacy improved due to the course, their attitudes toward assessment did not significantly change. They concluded that assessment literacy should be stressed in teacher education programs by considering students' perceptions and how these affects efficient practice.

Alkharusi et al. (2011) aimed to explore distinctions between teacher candidates

and teachers concerning their theoretical understanding, perceived proficiency in, and attitudes regarding assessment. In addition, the study aimed to investigate whether: (1) teaching internship influences teacher candidate's theoretical understanding, perceived proficiency, and attitudes regarding assessment; (2) year of experience influences teachers' theoretical understanding, perceived proficiency, and attitudes regarding assessment; (3) assessment course affects teachers' theoretical understanding, perceived proficiency, and attitudes regarding assessment. This study involved 512 participants, including 279 teacher candidates and 233 teachers. The Teacher Assessment Literacy Questionnaire (TALQ), developed by Plake and Impara in 1993, contains 35 items used as the data collection tool to measure participants' theoretical understanding. Firstly, the study found that, on average, the theoretical understanding of teachers was lower than that of teacher candidates but had more significant levels of perceived skillfulness and more favorable attitudes regarding assessment. Secondly, the study discovered that teacher candidates who had teaching internships while taking the assessment course had greater theoretical understanding, perceived proficiency, and more favorable attitudes regarding assessment than those who did not. Thirdly, the findings showed that teachers with fewer years of experience had higher theoretical understanding levels than more experienced teachers. Nevertheless, there were no significant differences in perceived proficiency between teachers with fewer or higher years of experience. Lastly, the study found that teachers who had taken an assessment course in undergraduate education had a greater theoretical understanding than those who did not. Nonetheless, the perceived proficiency of teachers who had taken the assessment course and those who had not showed no appreciable variations.

Akayuure (2021) conducted a study to examine the assessment literacy levels of high school mathematics teachers and their self-efficacy views concerning their assessment literacy. Additionally, the study investigated any correlations between the assessment literacy levels and self-efficacy views of high school mathematics teachers concerning their assessment literacy. The participants of the study are 96 high school mathematics teachers in Ghana. The researcher used the "Classroom Assessment Literacy Inventory" developed by (Mertler, 2003) to measure teachers' assessment literacy. According to the study results, teachers gave correct answers

to approximately 18 questions in the assessment literacy inventory consisting of 35 questions. Accordingly, high school mathematics teachers' assessment literacy levels were low. When the analyzes are examined in more detail, the most successful competency area of the teachers is 2 (developing assessment methods), while the most unsuccessful is 5 (developing a valid grading procedure). Moreover, the study found no variations in the assessment literacy levels of participants concerning their gender and year of teaching. Lastly, the study observed that the associations between the assessment literacy levels, and self-efficacy beliefs of the participants were primarily poor and not statistically meaningful, even after controlling for factors such as gender and year of teaching. According to the report, organizations that educate math teachers should emphasize assessment literacy improvement to guarantee consistent assessment procedures and high-quality math teaching in high schools.

Ergül and Çetin (2021) explored the assessment literacy levels of secondary school teachers. The researchers aimed to identify which areas of teacher competence were lacking and to explore the relationship between assessment literacy and year of teaching, subject area, and item analysis talents. The participants are 189 secondary school teachers with branches of mathematics, Turkish language, physical sciences, social studies, English language, and education of religion. In this study, the "Assessment Literacy Inventory" designed by Mertler and Campbell (2005) was used in its Turkish version, which was adapted by Bütüner et al. (2010). Participants responded to approximately 13 out of 30 questions accurately, showing that the teachers' assessment literacy levels were insufficient. The findings indicated that the teachers were most competent in the first standard (choosing assessment methods) with a mean score of 3.03, while they were least competent in the fifth standard (developing valid grading procedures) with a mean score of 1.39. There was also a reasonable positive correlation between teachers' ability to item analysis talents and their levels of assessment literacy. Moreover, teachers with more years of teaching demonstrated higher levels of assessment literacy than those with less. The researchers recommended conducting seminars or education on assessment, given teachers' low levels of assessment literacy. These programs can include theoretical and practical training to enhance teachers' understanding of assessment.

Bijsterbosch et al. (2019) research offered insights into how a teacher professional

development program contributed to attaining advanced levels of assessment literacy. The researchers employed Xu and Brown's"Teacher Assessment Literacy in Practice" framework. Eight participants took part in this research, which utilized qualitative methods. A content analysis of summative assessments was used to analyze teachers' practices to gauge their professional development in terms of assessment literacy. The researchers assessed potential alterations in the teachers' knowledge, skills, and beliefs by administering questionnaires and conducting interviews. Based on the study's results, it was revealed that contemplating educational goals and teachers' beliefs, participating in collaborative activities, and receiving feedback from peers were crucial factors in fostering higher levels of assessment literacy.

In the research carried out by Ayalon and Wilkie (2020), they explored the encounter of 60 pre-service mathematics teachers at the secondary level with a series of formative assessment activities, aiming to simulate the practices of in-service teachers in a school environment. The study's data comprises written responses from pre-service mathematics teachers, and as a result, qualitative data analysis was employed. They claimed that pre-service teachers can still learn through approximations of practice even if they have not yet had the chance to gain such assessment knowledge experientially. Additionally, they have mentioned that providing sample student solutions or answers and analyzing them contributes to the participants' assessment literacy.

2.3 Summary of the Literature Review

The assessment concerns a methodical approach to gathering, analyzing, and utilizing data linked to educational initiatives, with the primary goal of enhancing the growth and progression of students (Brown, 2008). This procedure implicates compiling information from various sources to acquire a comprehensive insight into the knowledge acquired by students and their ability to apply it based on their educational experiences (Chapelle et al., 2010). In this way, teachers can identify students' learning deficiencies, d etermine t heir i nterests a nd a bilities, a ssess their achievements, evaluate instruction effectiveness, and judge the curriculum (Gardner, 2012). Ultimately, the assessment results enhance upcoming learning experiences

and enrich students' understanding (Reynolds et al., 2008). Many studies show that teachers' competence in assessment is necessary to create an effective educational environment (Blömeke & Delaney, 2012; Cizek, 2010; Harrison et al., 2022; Suurtamm et al., 2010). For this reason, the literature emphasizes that it is important for teachers to have knowledge and skills in assessment and sees this as an area of competence that teachers should have (AFT et al., 1990; DeLuca et al., 2016; Popham, 2014; Price, 2005).

Teachers with the solid knowledge, skills, and understanding necessary to effectively design, implement, and evaluate assessments in an educational context are referred to as assessment-literate (Popham, 2004). Teachers who are proficient in assessment literacy should create high-quality assessments, demonstrate the knowledge and skills required to evaluate the validity and reliability of students' learning, consistently use standards and descriptions to evaluate and pass critique on students' work, and interpret and use assessment feedback and data to improve teaching and learning (Xu & Brown, 2016). While assessment literacy is considered an essential skill for teachers, numerous studies have highlighted the low assessment literacy levels among both pre-service and in-service teachers (Brookhart, 2001; Clark et al., 2022; DeLuca & Klinger, 2010; Mertler & Campbell, 2005; Popham, 2004; Stiggins , 1995; Ye, 2023). Research conducted with teacher candidates has shown that they lack sufficient knowledge in assessment and have various shortcomings in developing, implementing, and interpreting assessment approaches (DeLuca & Klinger, 2010; Kyttälä et al., 2022; Macken et al., 2020; Ogan-Bekiroglu & Suzuk, 2014; Volante & Fazio, 2007). Similarly, studies have indicated that even experienced teachers who have more opportunities to implement assessment in real classroom settings struggle to integrate assessment into their teaching effectively. They are inadequate in using appropriate assessment methods, analyzing and interpreting assessment results, or improving instruction based on assessment findings (Deneen & Hoo, 2023; Khadijeh & Amir, 2015; Koh, 2011; Pastore & Andrade, 2019).

Studies indicates that one of the reasons both teachers and teacher candidates have low levels of assessment literacy is the inadequacy of teacher education programs (DeLuca & Klinger, 2010; Deneen & Brown, 2016; Koh, 2011; Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk, 2014; Stiggins , 1995; Volante

& Fazio, 2007). McGee and Colby (2014) pointed out that the lack of emphasis on assessment within teacher preparation programs results in teacher candidates not acquiring the necessary knowledge and skills in this field.

According to the literature, there are deficiencies in both the theoretical and practical aspects of assessment education within teacher training programs (Pastore & Andrade, 2019). Teachers and teacher candidates have theoretical knowledge gaps in areas such as understanding different assessment methods (DeLuca & Klinger, 2010), knowing how to use appropriate assessment according to different classroom needs (Ladson-Billings, 2009), scoring, analyzing, and interpreting assessment results (Uludag & McDonough, 2022), or planning the next lesson based on assessment results (Sugianto, 2020). In practical terms, teachers often do not have enough opportunities to apply their assessment knowledge (Dehqan & Asadian Sorkhi, 2020; Essomba, 2010), there is a gap between theory and practice (DeLuca & Klinger, 2010), or they may not receive sufficient feedback on assessment during their internship experience from mentor teachers (Mohono-Mahlatsi & Van Tonder, 2006).

Although the literature highlights certain theoretical and practical deficiencies in teacher education programs regarding assessment, there is limited knowledge about the specific competency areas that teachers considered assessment-literate should possess. Therefore, this study aims to investigate the areas that need improvement in assessment education and recommendations of pre-service and in-service middle school mathematics teachers regarding teacher preparation programs for each of the competency areas presented by AFT et al. (1990).

CHAPTER 3

METHODOLOGY

The methodology of the study will be discussed in this chapter. It will cover ten subtopics, namely the design of the study, the context of the study, participants of the study, data collection tools, data collection procedures, data analysis, the trustworthiness of the study, ethical considerations of the study, assumptions of the study, and limitations of the study.

3.1 Design of the Study

The current study has three main aims. The first aim is to investigate the assessment literacy level of pre-service and in-service middle school mathematics teachers. The second and third aim of the study is to explore the perceptions of middle school mathematics pre-service and in-service teachers about the areas that need improvement in the teacher education program concerning assessment literacy and to investigate their suggestions for enhancing assessment literacy education provided in the teacher education program. Quantitative and qualitative research designs have been utilized to achieve these objectives. Quantitative research transforms occurrences and phenomena into quantifiable entities, rendering them capable of being observed, measured, and conveyed in numeric terms (Bloomfield & Fisher, 2019). A quantitative research design has been employed for the first research question, focusing on the assessment literacy levels of pre-service and in-service middle school mathematics teachers. The assessment literacy levels of the participants were determined by administering the assessment literacy inventory.

Qualitative research design was used to answer the second and third research

questions. Qualitative research is the study in which different qualitative data creation methods such as observation, interview, and document analysis are used, and a qualitative process is followed to reveal perspectives, perceptions, and events in a natural environment realistically and holistically (Yıldırım & Şimşek, 2013). In qualitative research, it is desired to obtain a much broader perspective rather than learning how much or how good the research subject is (Büyüköztürk et al., 2013). There is an effort to reach a deep perception of the investigated event or phenomenon (Morgan, 1996) and examine the quality of interactions, behaviors, circumstances, or resources in the research designed with the qualitative method (Fraenkel et al., 2011). Merriam and Tisdell (2015)) claim that by employing qualitative data-gathering techniques, including content analysis, observation, and interviews, qualitative research designs depict people's perspectives and experiences in their natural environments from a broad perspective. Accordingly, interviews were conducted to explore the perceptions of in-service and pre-service middle school mathematics teachers regarding the deficiencies of their teacher education program in terms of assessment literacy and their suggestions for enhancing the effectiveness of a teacher education program in terms of assessment literacy.

A case study, one of the qualitative research designs, was utilized in this study because the study aimed to comprehend unique situations within a specific context and the interactions occurring therein (Creswell, 2013; Patton, 2002). A case study involves conducting a thorough examination of a single person, group, or significant instance. Diverse data are gathered and analyzed to develop insights that can be applied to a particular case or to derive valuable general conclusions (Fraenkel et al., 2011). One of the main objectives of case studies is to understand a case along with its sequence of events and actions (Stake, 2006). The case in this study is pre-service and in-service middle school mathematics teachers who are studying in or graduated from the same teacher education program at the same university. Two issues were investigated in depth for both groups (teacher candidates and in-service teachers). These include participants' perceptions of shortcomings of assessment education offered by one of the teacher education programs and their suggestions for enhancing the effectiveness and quality of the teacher education program for developing assessment literacy of teacher candidates. Therefore, the use of a case study as a research design was

considered appropriate, aligning with the study's objectives and capable of addressing the second and third research questions.

3.2 Context of the Study

This research was carried out in a state university's Elementary Mathematics Education (EME) program in Ankara, Turkey. This four-year teacher education program is designed to prepare future mathematics teachers who will teach students in grades 5 to 8. In the first four semesters of the program, teacher candidates primarily take courses in mathematics, science, and general education. In the final four semesters, the focus shifts towards courses more specific to mathematics teaching. The teacher education program offered by the university where the research was conducted was updated in 2018. Accordingly, all in-service teachers received an education based on the program before 2018. On the other hand, all pre-service teachers received an education based on the program implemented after 2018. In both programs, some of the courses that teacher candidates took related to assessment include an assessment course, statistics courses, and two teaching practice courses (where they can apply the theoretical knowledge they have acquired).

All participants in this study were either senior-year students or graduates of the EME program of the university where the data was collected. The participants have taken several courses involving assessment during their undergraduate education, such as measurement and assessment, statistics, and teaching practice (internship). Among them, only one compulsory course was solely on assessment in mathematics education. This course, prior to 2018, was offered in the fourth semester under the code ELE225 and titled "Measurement and Assessment," whereas in the curriculum after 2018, it was offered in the fifth semester under the code MSE305 and titled "Assessment of Learning in Science and Mathematics." Both courses center around creating and utilizing classroom assessments to evaluate students' understanding in alignment with learning goals, interpretation of test results, fundamental statistics, and reporting. To elaborate further, the goals of both courses include gaining a comprehensive understanding of fundamental concepts in assessment and measurement, recognizing the significance of measurement and assessment in the

teaching process, identifying instructional objectives as desired learning outcomes, comprehending the factors involved in creating effective assessment instruments for classroom evaluation, and creating a variety of assessment materials for classroom use (Anonymous, Academic Catalog, 2022). In both the program before 2018 and the program after 2018, the outcomes and objectives of the assessment course are very similar. Therefore, the appendix section includes the curriculum of the assessment course in the program updated after 2018 (see Appendix A).

At the university where the study was conducted, two statistics courses are compulsory for students studying in the elementary mathematics teaching program. The name of the statistics course is "Introduction to Probability and Statistics," and it is divided into two parts, with one being offered in the third semester and the other immediately following it in the fourth semester. While the codes of the courses before 2018 were STAT201 and STAT202, respectively, they became STAT203 and STAT204 on and after 2018. However, the content and duration of the courses did not change. The course "Introduction to Probability and Statistics I" includes the following subtopics in statistics: descriptive statistics, probability, combinatorial methods, conditional probability, random variables, univariate and bivariate distributions, expectation, variance, covariance, correlation, some useful distributions, central limit theorem, and estimation. The course "Introduction to Probability and Statistics II" includes the following subtopics in statistics: Transformations of random variables, generating functions, conditional expectation, limit theorems, central limit theorem, and limiting distributions (Anonymous, Academic Catalog, 2015 & 2020).

The other courses that involve assessment are teaching practice (internship) courses, which are offered in the two semesters of the final year of the teacher education program. In the curriculum before 2018, the first course was coded as ELE419 and titled "School Experience," whereas, in the updated curriculum from 2018 onwards, the course is coded as MSE409 and titled "Practice Teaching I." The description of the course with the code ELE419 includes classroom observation, which encompasses various aspects such as the structure and control of the school environment, daily routines within the school, group activities, a typical day for teachers and students, collaboration between the school and families, observation of both core

and non-core subjects, examination of school-related issues, diverse teaching and learning activities, and the assessment of materials and written resources. The description of the course with the code MSE409 involves field experience and practice teaching, which involve a range of activities such as observing classrooms, engaging in micro-teaching, developing instructional activities and materials, planning and preparing for teaching, managing the classroom, conducting measurement and assessment, and reflecting on the teaching experience. The main difference between these two courses is that the updated curriculum from 2018 onwards emphasizes more on practical application rather than observation (Anonymous, Academic Catalog, 2015 & 2020).

The predecessor of Practice Teaching II, which is the continuation of Practice Teaching I, had the code ELE420 and was titled "Practice Teaching in Elementary Education" before 2018. Both ELE420 and MSE410 have the same content. In both courses, the aim is to provide field experience and practice teaching, which involves a range of activities such as observing classrooms, engaging in micro-teaching, developing instructional activities and materials, planning and preparing for teaching, managing the classroom, conducting measurement and assessment, and reflecting on the teaching experience (Anonymous, Academic Catalog, 2015 & 2020).

3.3 Participants of the Study

The study's target population is the pre-service and in-service middle school mathematics teachers who have taken the teacher education courses including the assessment course at the same university. At the university where the study was conducted, the target population for pre-service middle school mathematics teachers in the research is approximately 100 individuals. The target population for in-service middle school mathematics teachers consists of those who have graduated within the last 10 years. Since reaching every target group member was neither feasible nor practicable, non-random sampling methods were used in this study in two stages according to the study's aims. The process of selecting participants is described in the following sections.

3.3.1 First Stage of the Sample Selection

The first aim of the study is to investigate the assessment literacy levels of pre-service and in-service middle school mathematics teachers. To achieve this goal, it was intended to administer the assessment literacy inventory to as many participants as possible, aiming to gain a comprehensive perspective on the assessment literacy levels of pre-service and in-service middle school mathematics teachers. For this reason, the researcher tried to reach as many pre-service and in-service teachers as possible by using convenience and purposive sampling methods, with the support of face-to-face, mail, electronic communication tools, and key persons, such as instructors. When choosing a random or systematic non-random sample is challenging, researchers can use the convenience sampling method to select participants who are suitable for the study in terms of cash, time, and place (Creswell, 2013; Fraenkel et al., 2011). In purposive sampling, researchers choose individuals who will supply the required information based on specific criteria to comprehend a phenomenon (Creswell, 2013; Fraenkel et al., 2011). Merriam (2009) claims that before beginning a purposive sampling, the selection criteria must be determined. The selection criterion for pre-service teachers was that they should have taken the assessment course. The criterion for selecting in-service teachers was that they had recently graduated from teacher education programs, aiming to facilitate a more accurate and comprehensive recollection of their assessment experiences. Additionally, it was ensured that they were graduates of the Elementary Mathematics Education (EME) program at the same university where the data of pre-service teachers were collected. The reason for choosing pre-service and in-service middle school mathematics teachers from the same university is the assumption that both groups have similar undergraduate education experiences. The researcher relied on this assumption by examining the academic catalogs of courses from different years offered by the university where the study was conducted. In her analysis, the researcher observed that although there were some changes in the course contents offered in academic catalogs over the years, they remained similar in terms of learning objectives and topics (more detailed information is provided in the "context of the study" section). Thus, it aims to enrich the study's data with the fresh experiences of the teacher candidates and the teachers' teaching experiences. Furthermore, examining situations where the

experience differences between pre-service and in-service teachers who graduated from the same university are more pronounced can provide a better understanding of the program's effects. For instance, the opinions of pre-service and in-service teachers with different grade levels or varying teaching experiences can be more readily compared.

Snowball sampling was one of the purposive sampling techniques employed in the study's first sample selection stage. In the snowball sampling method, the participants direct the researcher to the people who can contribute to the research so that the study participants grow and multiply like a snowball (Merriam, 2009; Patton, 2002). In this study, the initial participants were selected from the city where the researcher was living due to the convenience and easy accessibility for them. Although approximately 70% of the participants were initially selected, it became apparent that the number of participants needed to be increased. Recognizing this, the researcher asked these initial participants if they knew of any other individuals who could join the study. Following the recommendations of these initial participants, the researcher reached out to several more potential participants through phone calls and emails. Each participant suggested potential participants for the study. Consequently, the remaining participants were selected based on the suggestions of the initial participants, forming the composition of the study's participants. As a result, an assessment literacy inventory was administered to 34 pre-service and 41 in-service teachers using both convenience and snowball sampling methods. The characteristics of the participants who participated in the first stage of the study are given in detail in the table below.

Table 3.1: Major Characteristics of the Participants Who Participated in the First Stage of the Study

Participants	**Assessment Literacy Level**	**Gender**		**Frequency**	**Percent**
		Female	**Male**		
Pre-service	15.94	31	3	34	45
In-service	15.73	31	10	41	55
Total	31.67	62	13	75	100

3.3.2 Second Stage of Sample Selection

The applied assessment literacy inventory of 34 pre-service and 41 in-service teachers in the first stage was analyzed based on assessment literacy competencies stated by AFT et al. (1990), and the assessment literacy levels of the participants were examined. While applying for the Assessment Literacy Inventory, the participants were also asked whether they volunteered to participate in the interviews for the next stage. 18 of the 34 pre-service teachers and 21 of the 41 in-service teachers stated that they volunteered for the interviews. The participants to be interviewed were purposefully selected from among these voluntary participants. The purposeful sampling method (Patton, 2002) was employed to select participants who would produce rich data for the study and be more beneficial in understanding the phenomenon. Maximum variation sampling is one of the methods used for purposive sampling that entails locating and looking for participants who reflect the greatest variety of study-relevant features (Merriam, 2009). Accordingly, using maximum variation sampling, 5 out of 18 pre-service teachers and 5 out of 21 in-service teachers were selected for interviews according to their various characteristics. To promote variety in the selection of individuals interviewed, the researcher established the following criteria:

- The middle school mathematics pre-service teachers to be interviewed were selected according to their, grade level, GPA, assessment literacy levels and whether they took the teaching practice courses.

  Grade levels: Pre-service teachers participating in the study were tried to be selected in a variety of ways according to whether they were in the 3rd or 4th grade.

  GPA: Teacher candidates were selected to represent the performance levels of "3.51 – 4.00," "2.51 – 3.00," "2.50 – 2.00," and "less than 2.00" based on their grade point averages.

  The teaching practice courses: According to the literature, applying the theoretical knowledge learned for assessment, that is, gaining experience, is an essential factor affecting the assessment literacy levels of pre-service and in-service teachers, as well as theoretical knowledge (DeLuca & Klinger, 2010;

Mertler, 2009; Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk, 2014; Volante & Fazio, 2007). For this reason, while selecting pre-service teachers to interview, attention was paid to whether they had teaching practice experience or not.

Assessment literacy levels: Participants with low, medium, and high assessment literacy levels were tried to be selected from the pre-service teachers who volunteered for the interview part of the study. Thus, it aimed to collect richer data regarding the deficiencies of undergraduate education, and suggestions for improving undergraduate education in this regard.

For these purposes, 5 pre-service teachers with different characteristics were selected for the interviews. Middle school pre-service mathematics teachers were coded with P codes, for example, P1, P2. More detailed characteristics of pre-service teachers are given in Table 3.2 below:

Table 3.2: Characteristics of Pre-service Teachers Who Participated in the Interviews

Code for Participant	**Gender**	**Grade Level**	**GPA**	**Taking Teaching Practice Courses**	**Assessment Literacy Score (out of 30)**
P1	Male	3	2.51 - 3.00	Teaching Practice I	18
P2	Female	4	2.51 - 3.00	Teaching Practice I and II	14
P3	Female	4	3.01 - 3.50	Teaching Practice I and II	20
P4	Female	4	3.01 - 3.50	Teaching Practice I and II	16
P5	Male	4	2.00 - 2.50	Teaching Practice I and II	13

- The middle school in-service mathematics teachers to be interviewed were selected according to their assessment literacy levels, years of teaching experience, the type of school they work at, and their educational status.

  The years of teaching experience: In-service teachers may have had different experiences regarding assessment according to their years of teaching experience. Teachers with 0-10 years of teaching experience were selected in line with that.

  The type of school: Assuming that the experiences obtained in private and public schools may differ, in-service teachers working in public and private schools were selected.

  The educational status: Assuming that in-service teachers' assessment literacy may be affected by the level of education, teachers with master's or doctorate degrees were also selected.

  Assessment literacy levels: The same reasons mentioned above with pre-service teachers also apply to in-service teachers.

For these purposes, 5 in-service teachers with different characteristics were selected for the interviews. More detailed characteristics of in-service teachers are given in Table 3.3 below:

Table 3.3: Characteristics of In-service Teachers Who Participated in the Interviews

Code for Participant	**Gender**	**Years of Experience**	**School of Employment**	**The Educational Status**	**Assessment Literacy Score (out of 30)**
T1	Male	0 – 2	State School	M.S. in Mathematics Education (ongoing)	18
T2	Female	3 – 5	State School	PhD. in Mathematics Education (ongoing)	20
T3	Male	0 – 2	Private School	B.S. in Mathematics Education	13
T4	Male	3 – 5	State School	B.S. in Mathematics Education	15
T5	Female	0 – 2	State School	B.S. in Mathematics Education	14

3.4 Data Collection Tools

Two instruments were used to collect data from the participants to address the aims of the study. In order to collect the necessary data to explore the assessment literacy levels of pre-service and in-service middle school mathematics teachers (the first aim), the Assessment Literacy Inventory (Mertler & Campbell, 2005) was

applied to the participants. To collect data on pre-service and in-service middle school mathematics teachers' thoughts about and evaluation of the teacher education program in terms of developing their assessment literacy semi-structured interviews were used. Details of these two data collection tools are given in the next sections.

3.4.1 Assessment Literacy Inventory

Mertler and Campbell (2005) developed the "Assessment Literacy Inventory" to investigate teachers' and teacher candidates' assessment literacy levels. They developed the assessment literacy inventory based on the seven assessment competency areas introduced by the AFT, NEA, and NCME (1990). These standards are 1) teachers need to possess the competence to select proper assessment techniques, tools, and methods for teaching decisions; 2) teachers need to possess the competence to generate proper assessment techniques, tools, and methods for teaching decisions; 3) Teachers need to possess the competence to administer, score, and interpret the assessment results; 4) Teachers need to possess the competence to use assessment results for making decisions related to students, instructional planning, curriculum development, and school improvement; 5) Teachers need to possess the competence to create accurate techniques for grading students based on their assessments; 6) Teachers need to possess the competence to explain assessment findings to students, colleagues, and parents; and 7) Teachers need to possess the competence to identify assessment information that is unethical, illegal, or otherwise inappropriate. The assessment literacy inventory consists of five scenarios and seven questions about each scenario (35 items in total). Mertler and Campbell (2005) created the measurement tool with a two-stage pilot test. In 2002, in the first pilot test, they applied the inventory to 152 pre-service teachers. The first pilot test's KR20 (rKR20) reliability was .75, the mean item difficulty was .64, and the mean item discrimination was .32. While these values indicate that the generated inventory performs quite well, the inventory has been revised to enhance its other psychometric properties. After the revisions, the second pilot test was conducted in 2004 and applied to 249 teacher candidates. The second pilot test's KR20 (rKR20) reliability was .74, the mean item difficulty was .681, and the mean item discrimination was .313. In the literature, different researchers have suggested different values regarding

the minimum value required for a measurement tool to be considered reliable. The results of assessment literacy developed by Mertler and Campbell (2005) show that the constructed inventory's psychometric qualities significantly support its usage as a valid indicator of teachers' assessment literacy.

Bütüner et al. (2010) adapted the Assessment Literacy Inventory to Turkish. The Turkish version of the inventory consisted of 30 items because the items belonging to the seventh competence area of the actual inventory (recognizing immoral and illegal practices) do not match the performance indicators representing the competence areas determined by the Ministry of National Education in Turkey. Therefore, Bütüner et al. (2010) did not include the part of recognizing immoral and illegal practices in the Turkish version of the inventory. The competence areas of the items in the inventory are shown in Table 3.4.

Table 3.4: Competence Areas and the Corresponding Items for the Adapted Version of the Assessment Literacy Inventory (Bütüner et al., 2010)

The Standards for Teacher Competence in the Educational Assessment of Students (1990)	**Items**
Standard 1: Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.	1, 7, 13, 19, 25
Standard 2: Teachers should be skilled in developing assessment methods appropriate for instructional decisions.	2, 8, 14, 20, 26
Standard 3: The teacher should be skilled in administering, scoring and interpreting the results of both externally produced and teacher-produced assessment methods.	3, 9, 15, 21, 30
Standard 4: Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.	4, 10, 16, 22, 28
Standard 5: Teachers should be skilled in developing valid pupil grading procedures which use pupil assessments.	5, 11, 17, 23, 29
Standard 6: Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.	6, 12, 18, 24, 27

The six proficiency areas covered by the original inventory were reduced to four proficiency areas determined by the Ministry of National Education by Bütüner et al. (2010). Table 3.5 shows which competence areas in the original inventory correspond to those in Turkey.

Table 3.5: Comparison of the Assessment Competency Areas in the Original Inventory and the Competencies Stipulated by the Ministry of National Education in Turkey (Bütüner et al., 2010)

Assessment Competency Areas in Turkey	**Corresponding Competency Area in the Original Inventory**
Competency Area 1: Determining assessment methods and techniques	Standard 1
Competency Area 2: Measuring students' learning by using various assessment methods	Standard 2 and 5
Competency Area 3: Analyzing and interpreting data, giving feedback on student progress and learning	Standard 3 and 6
Competency Area 4: Reviewing the learning teaching process according to the results	Standard 4

Bütüner et al. (2010), applied the inventory they adapted into Turkish to 260 pre-service teachers. The researchers analyzed the items' difficulty discrimination values and the Kuder-Richardson Reliability Coefficient (KR-20) of the inventory. According to the researchers' analysis, the KR-20 reliability value was calculated as 0.86. In addition, the mean item difficulty level was calculated as 0.644, and the mean item discrimination level was 0.486. Based on these values, the inventory can be accepted as a reliable measurement tool (Chase, 1999; Kehoe, 1994; Nitko, 2001). Therefore, in this study, the inventory adapted into Turkish by Bütüner et al. (2010) was used to determine the assessment literacy levels of pre-service and in-service teachers. The "Assessment-Evaluation Literacy Inventory" used in the study is presented in Appendix B. The inventory was used after permissions were received from Mertler and Campbell, who developed the inventory, and Bütüner, who adapted the inventory to Turkish.

3.4.2 Interview Protocol

One of the most used methods to collect data in qualitative research is interviews (Creswell, 2013; Fraenkel et al., 2011). Through interviews, researchers can conduct in-depth research and better understand people's thoughts, feelings, experiences, and perceptions (Patton, 2002). For this reason, after the assessment literacy inventory was applied, semi-structured interviews were conducted with 5 pre-service teachers and 5 pre-service teachers selected from the participants to whom the inventory was applied. The information gathered through these interviews was then used to examine the deficiencies of the university's teacher education program for assessment literacy and participants' suggestions for better quality undergraduate education on assessment literacy.

The researcher created open-ended questions for semi-structured interviews. The interview questions were formed based on the six competency areas used while creating the assessment literacy inventory. In creating the interview questions, the related literature was reviewed, and the results of other studies were also considered. For instance, studies focusing on assessment literacy with teacher candidates have revealed a lack of practical application in teacher education (Beziat & Coleman, 2015; DeLuca & Klinger, 2010; Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk, 2014; Volante & Fazio, 2007). Therefore, one of the considerations taken into account when preparing interview questions was the assessment practices of the teacher education program. After the interview questions were composed, they were refined and finalized by taking an expert's opinion. Interview questions begin with questions about participants' thoughts about the role and importance of the assessment. The following questions consist of two main sections: deficiencies of undergraduate education for each competency area and suggestions for better quality undergraduate education. Under these three main sections, follow-up questions were added according to the participants' answers.

Before conducting interviews with actual participants, the prepared interview questions were checked for their suitability and comprehensibility by conducting pilot interviews. At the beginning of the interviews, the purpose of the study and the interview was explained to the participants. The first pilot interview was conducted

with two people, one pre-service and one in-service teacher. After the first pilot interview, some interview questions were slightly modified. For example, in the first pilot study, participants had difficulty fully grasping the content from the names of the competency areas, and the researcher recognized the need for more detailed explanations. Therefore, in the second pilot study conducted after the first one, the researcher provided a brief explanation followed by examples for each competency area, aiming to offer participants a more comprehensive understanding. As a result, it was noticed that participants better comprehended and recalled the discussed topics. The second pilot study was conducted with two participants, one being pre-service and the other in-service. After the second pilot interview, the interview questions took their final form with the guidance of the expert. For instance, one of the changes made under expert guidance was adding the question "What kind of experience can other public or private institutions and organizations provide to assist your development in measurement?" to the interview questions. The final version of the interview protocol can be seen in Appendix C and some of the interview questions are given below.

Some of the interview questions asked to pre-service teachers are:

- How did the assessment/method/ internship/statistic courses you took in your undergraduate education fell short in choosing appropriate assessment methods?
- How did the teacher education program fall short in your choosing appropriate assessment methods?
- Was the number or content of the elective courses offered in the teacher education program for assessment sufficient? What would you suggest doing about this issue?

Some of the interview questions asked to in-service teachers are:

- When you became a teacher, was the information you gained from your undergraduate education sufficient to choose appropriate assessment methods?
- What were your concerns in choosing appropriate assessment methods before you started teaching?

- If you have experiences that help you develop in terms of assessment after becoming a teacher, what would be your suggestions for integrating these experiences into teacher education programs?

3.5 Data Collection Process

The data of this study was collected from pre-service and in-service middle school mathematics teachers. The data of the assessment literacy inventory in this study were collected at the end of the spring semester of 2021-2022. After collecting the data from the assessment literacy inventory, interview questions were prepared. After the interview questions were prepared, the researcher made a pilot study to review the interview questions' trustworthiness, suitability, and comprehensibility. The pilot study was implemented at the beginning of the summer of the 2021-2022 academic year. As a result of the pilot study, some adjustments were made to the interview questions, and the questions were finalized. After the pilot study, the interviews were conducted in the fall of 2022-2023. Before data collection, necessary permissions were obtained from the Middle East Technical University (METU) Applied Ethics Research Center (see Appendix F).

Before starting data collection, the researcher informed the participants about the purpose of the research, assessment literacy, data collection tools, and data collection process. They were also informed that participation in the study was voluntary and that they would withdraw at any point throughout the data collection process. This information was also conveyed to the participants with a written informed consent form, and participants who volunteered for the study signed this consent form (see Appendix E).

The data collection began by administering the assessment literacy inventory to pre-service teachers and in-service teachers. While the researcher communicated the assessment literacy inventory and consent form to some participants face-to-face, she sent them to others via e-mail and electronic communication tools. Before the university entered the summer break, the researcher obtained permission from the faculty members at the education department and distributed the data collection

instrument in person to the pre-service teachers. After the pre-service teachers, while the in-service teachers were on summer vacation, the researcher provided the inventory online to the teachers living in different cities. Participants were not given a time limit to complete the inventory in both conditions. Participants were asked to submit the inventory to the researcher approximately one week after receiving it. Data were collected from pre-service and in-service middle school mathematics teachers within about one month.

Interviews with pre-service and in-service teachers were held through the Zoom platform, one of the online meeting tools. The interviews with the participants were recorded with their consent, and the researcher also took notes during the interview. Interviews lasted approximately 90-100 minutes, and five or ten-minute breaks were given whenever the participants wanted, depending on their needs. In addition, while interviewing the participants, the researcher ensured that no one else was in the room so that the interview environment was quiet and the conversations continued without interruptions.

3.6 Data Analysis

Two separate sets of data were analyzed to fulfill the study's goals. These two datasets are the participants' responses to the assessment literacy inventory and transcripts of the interviews. First, the responses of pre-service and in-service teachers to the assessment literacy inventory were analyzed and their assessment literacy levels for the six competence areas were determined. Then, the transcripts of the interviews were analyzed and coded considering the assessment literacy levels of the participants, and then they were divided into categories. Yamtim and Wongwanich (2014), who aimed to investigate the classroom assessment literacy levels of elementary school teachers, categorized assessment literacy scores into three levels as follows: weak or requiring advance (below 60%), medium (60-79%), and well (80% and above). In this study, participants' assessment literacy levels were also classified, considering the study by Yamtim and Wongwanich (2014). Detailed information about the analysis of these two data sets is given in the following sections.

3.6.1 Analysis of Data from the Assessment Literacy Inventory

The quantitative data obtained from the assessment literacy inventory were analyzed using the SPSS program. Descriptive statistics such as mean, median, standard deviation, skewness, kurtosis, and the maximum and minimum values of the data were utilized to assess the levels of assessment literacy to analyze the essential traits of the sample. These descriptive statistics were examined separately for both pre-service and in-service teachers and conducted separately for each competency area. According to the competency areas, the frequencies of correct and incorrect answers to the items were determined. Since the assessment literacy inventory consists of multiple-choice questions, each question has only one correct answer. A correct answer was graded as 1 point and an incorrect answer was graded as 0 points. In order to determine whether the obtained data set showed a normal distribution, the given Skewness-Kurtosis values and The Shapiro-Wilk Test values were examined. According to the results obtained, the values were normal.

A t-test is a statistical method used to evaluate the means of two groups for comparison. The independent t-test is one type of t-test, applied when the two compared groups are not dependent on each other (Kim, 2015). In this study, an independent t-test was also utilized to determine whether there was a significant difference in the assessment literacy levels between pre-service and in-service middle school mathematics teachers. "Participant type (pre-service or in-service)" was taken as the categorical variable, and "assessment literacy level" was used as the continuous variable.

3.6.2 Analysis of Data from Interview Transcripts

Qualitative research aims to discover and understand the concepts or phenomena to be investigated. The researcher infers broad themes from specifics in the data analysis and then interprets the significance of the data (Creswell, 2013; Patton, 2002). In other words, qualitative data analysis seeks to create meaning about the data under study and explain what is represented in the data set through categorization and interpretation (Miles et al., 2020). Merriam (2009) argues that since qualitative data

analysts concentrate on the text of interviews and transcripts when analyzing data, all qualitative data analysis is content analysis. Content analysis is a methodical and consistent way of analyzing a piece of text by breaking it down into smaller parts and categorizing it based on specific predetermined criteria (Büyüköztürk et al., 2013). In the current study, the researcher concentrated on the thoughts and perceptions of pre-service and in-service middle school mathematics teachers about the assessment education that they received in their undergraduate education. Therefore, content analysis was used for a semi-structured interview in this study.

The answers given by the pre-service and in-service middle school mathematics teachers to the interview questions constituted the data in the main part of the study. After the interviews were done through Zoom meetings, the researcher transcribed all the audio interviews. Then, the researcher sent interview transcripts to three participants and asked them to review them. This allowed the researcher to ensure the accuracy of the collected data and provided participants with the opportunity to confirm that their words and experiences were accurately represented in the research. During this period, the study participants did not provide any feedback regarding inconsistencies. After, the transcripts were coded using the open coding method in the initial phase of the study. According to (Strauss & Corbin, 2004, p. 303), open coding is "the process of breaking down, examining, comparing, conceptualizing and categorizing data." During the open coding process, occurrences are assigned labels and brought together through continuous comparison in order to create categories and characteristics (Babchuk, 1996). During open coding, the researcher also coded line by line for each competency area. The data is coded in more detail with line-by-line coding, thus obtaining a more reliable analysis (Charmaz, 2008; Miles et al., 2020). In the current study, the obtained data was also read line by line and initially segmented. Subsequently, these segments were analyzed, examined, and subjected to comparisons in order to perform labeling and categorization. Accordingly, the researcher initially segmented the collected data according to each competency area. Subsequently, the deficiencies in the assessment literacy of the teacher education program and the participants' suggestions were separately analyzed for each competency area, leading to the creation of labels. Additionally, the researcher incorporated insights from the literature while forming

these labels. For instance, upon reviewing the literature, it was determined that the teacher education program had deficiencies both in theoretical and practical aspects of assessment literacy (DeLuca & Klinger, 2010; Pastore & Andrade, 2019; Siegel & Wissehr, 2011; Volante & Fazio, 2007). Consequently, while examining the deficiencies of the teacher education program for each competency area, the researcher took note of both theoretical and practical shortcomings. In doing so, the researcher meticulously labeled both deficiencies and suggestions line by line for each competency area, differentiating between practical and theoretical aspects. Following this, the theoretical and practical deficiencies were further analyzed and grouped within themselves. For example, codes related to managing, scoring, and interpreting assessment results were labeled for the third competency area. After labeling both theoretical and practical deficiencies for all competency areas, the researcher proceeded to the coding phase for the collected data. In the subsequent stage, the researcher categorized the obtained codes into groups. For instance, within the practical deficiencies of the fourth competency area, one code was "not using assessment results when preparing lesson plans," while another was "not preparing lesson plans that build upon each other in a logical sequence." After consultations with thesis advisor, these two codes were merged into the category "insufficient practice in planning the subsequent lesson according to the assessment results." In summary, the researcher meticulously analyzed each code line by line, leading to the creation of labels; these labels were then used to form codes, which in turn were grouped into categories.

Moreover, the researcher categorized the participants' suggestions for providing higher quality education in terms of assessment of undergraduate education into the framework presented by Juanjuan and Mohd Yusoff (2022). They aimed to examine the common characteristics found in successful programs that enhance teachers' assessment literacy in their study. Three themes were used to group the improvement programs' effectiveness by researchers: a) "What" is about the teacher's assessment literacy advancement program's content, b) "How" is about the teacher's assessment literacy advancement program's method, and c) "Why" is about the underlying orientation of the teacher's assessment literacy advancement program. For what (content), the researchers stated that sufficient assessment literacy

advancement programs should be customized to teachers' specific assessment training requirements, considering the local context and the day-to-day assessment practices in schools. How (method) fosters collaboration, promotes active participation and involvement, and offers chances for ongoing professional development. Lastly, why (orientation) advocates that the advancement programs should concentrate on assisting instructors in rethinking their position as assessors and aiding them in becoming reflective practitioners.

After all the data of the 10 participants' interview questions were coded, the researcher consulted thesis advisor and then rechecked the data codes. During the coding, the researcher aimed to ensure the validity of the codes she gave by asking the participants about the parts of the data analysis that she needed clarification on. Subsequently, a second coder, an experienced doctoral student specializing in mathematics education and qualitative studies, analyzed two sets of data, one from a pre-service participant and the other from an in-service participant. This step was taken to ensure the validity of the data analysis, which was based on semi-structured interviews. The codes from the second coder and the researcher were found to be approximately 80% in agreement. Although the codes of the second coder and the researcher were compatible, a small part of the codes was revised by reaching a full consensus after discussing the codes and categories with the second coder. For example, some codes that were very lengthy and therefore difficult to understand were condensed into shorter and more concise forms. Moreover, an expert in the field of measurement and evaluation examined a portion of the final analysis to check if/how the codes appropriately represent the issues on pre-service teacher education in assessment. Then the expert's suggestions were used to clarify the categories and remove the ambiguities in the coding scheme. The codes, sub-categories, and main categories are listed in the findings section. The findings section shows sample quotes from the teachers' responses that were deemed helpful in representing the results of the analysis.

3.7 The Trustworthiness of the Study

Regardless of quantitative or qualitative research, the research should be valid and reliable (Merriam, 2009). However, due to the nature of qualitative research, data analysis depends on the researcher's viewpoint (Fraenkel et al., 2011). Therefore, the concepts of validity and reliability in qualitative research are handled differently from quantitative studies, using different terminology (Miles et al., 2020). In qualitative research, validity and reliability are considered measures of trustworthiness, and four criteria - credibility, transferability, dependability, and confirmability - are used to enhance their credibility. Accordingly, in qualitative studies, credibility instead of internal validity, transferability instead of external validity, dependability instead of reliability, and confirmability instead of objectivity are used (Lincoln & Guba, 1985). This section covered four aspects of the study: credibility, transferability, dependability, and confirmability.

Credibility, which refers to internal validity, is the first factor to consider when assuring the trustworthiness of qualitative research. The concept of credibility pertains to whether or not the results of a research study accurately reflect reality (Merriam & Tisdell, 2015). Researchers must provide a detailed and accurate depiction of the studied situation and consistently interpret their results to establish credibility (Shenton, 2004). According to Merriam (2009), various approaches can be employed to enhance a study's credibility, such as peer examination, triangulation, adequate engagement in data collection, the researcher's position (reflexivity), and member checks. To strengthen the credibility of the current study, member checks, the researcher's position (reflexivity) and peer examination were all used.

The researcher's position involves a critical reflection on oneself as the researcher and the instrument used in the study. Researchers must articulate their biases, tendencies, and presumptions regarding the research they are about to undertake (Merriam & Tisdell, 2015). A potential limitation of this study was that the researcher and participants had completed the same teacher education program (Elementary Mathematics Education). The researcher acknowledged that she may have interpreted the results biasedly due to her familiarity with the participant's course experiences. Nonetheless, the fact that the researcher and participants shared the same context

also presented certain benefits. For instance, it allowed the researcher to develop semi-structured interviews that were well-suited to the participants' experiences and to identify overarching trends among them, which was one of the goals of the study. Another way of enhancing the credibility of a study is through peer review. This approach presents the study to at least one colleague or committee member knowledgeable about the topic and research methodology to obtain their opinion and approval. The reliability criterion is fulfilled by incorporating their feedback on the research ((Merriam & Tisdell, 2015). The researcher received feedback, comments, and suggestions from their thesis supervisor and committee members on the data collection tools, interviews, and findings to implement the peer review strategy. These inputs were considered in revising the final version of the study.

Transferability, also called external validity, is the second factor for guaranteeing the trustworthiness of qualitative research. The ability to generalize the results of a study to situations in similar participants and settings is called transferability (Lincoln & Guba, 1985). However, given that the outcomes of a qualitative study are tailored to a limited number of specific environments and individuals, it is not feasible to demonstrate the generalizability of the findings and conclusions to other populations and situations (Shenton, 2004). According to Merriam (2009), presenting rich and detailed descriptions is a procedure to enhance qualitative research studies' transferability. Therefore, in this study, the researcher aimed to offer ample and comprehensive explanations of the study's context and phenomenon as the participants, data collection process, data analysis process, and the current study's findings to enable readers to apply the study's outcomes to other settings to ensure the transferability of the findings.

Dependability, which refers to reliability in qualitative research, is the third factor for confirming trustworthiness. The concept of dependability is whether the research procedure is uniform among different researchers and research methodologies, reasonably consistent across time, and coherent (Gay et al., 2006; Miles et al., 2020). Merriam (2009) identifies four methods to improve the dependability of qualitative research studies, which are: (1) peer examination, (2) audit trail, (3) triangulation, and (4) researcher's position. This study used peer assessment and researcher's position methods to provide consistency in the results. To ensure the utilization of the audit

trail strategy, the procedures for gathering, interpreting, and analyzing data were thoroughly elucidated. In the section on credibility, the researcher explained how she employed these different methods.

Confirmability, or objectivity, is the final factor for assuring trustworthiness in qualitative research. To ensure objectivity in a study, minimizing the influence of the researcher's personal biases and preferences is essential. This involves taking steps to ensure that the findings and conclusions are based on the experiences and perspectives of the participants rather than those of the researcher (Gay et al., 2006; Shenton, 2004). To assure the objectivity of the present study, a comprehensive account of the entire research procedure and the researcher's role in it was previously presented.

3.8 Ethical Considerations of the Study

It is crucial to comply with ethical standards during research to protect the participants' well-being and maintain the study's integrity (Gregory, 2003). According to Fraenkel et al. (2011), three key issues that researchers must prioritize for an ethical study: (1) safeguarding study participants from harm, (2) preserving the confidentiality of research data, and (3) addressing the potential for the deception of participants. As there is no deception of the participants for this study, the issues of protecting the participants from harm and ensuring the confidentiality of the research data are discussed.

Before starting data collection, the researcher informed the participants about the purpose of the research, assessment literacy, data collection tools, and data collection process. They were also informed that participation in the study was voluntary and that they would withdraw at any point throughout the data collection process. This information was also conveyed to the participants with a written informed consent form, and participants who volunteered for the study signed this consent form (see Appendix E). Initially, the ethics committee at the university where the participants were registered granted the authorization to fulfill the current research to apply inventory and interviews (see Appendix F). The first ethical issue was ensuring that participants were not harmed physically or psychologically. At

the start of the study, the researcher conveyed verbal and written details to the participants about the study's objectives and the procedures that would be employed and highlighted that participation was entirely voluntary and that they could opt out without any repercussions. The participants were provided with written information about the study's objectives and methods, and a consent form was given to them, which they signed before the interviews began. Additionally, before both data collection processes, the participants were provided with a consent form containing this information, which they had to sign before the interviews could start. Before recording the interviews with the online meeting tool, permission was requested from one participant again. The participants were assured that their responses to the questions were not right or wrong, and the researcher refrained from passing any judgment during the interview to establish a comfortable setting for the participants.

The second ethical issue considered was preserving the confidentiality of research data. The researcher conveyed to the participants that the researcher would not share the collected data with anyone other than the researcher and the dissertation advisor and would keep the participants' answers confidential. Moreover, the study's advisor did not know the names of the participants, although he had access to the collected data. To prevent any direct correlation between the data and the participants' names when the study's conclusions were presented, the researcher changed real names with nicknames such as T1 and P1. Pre-service teachers were coded with the code P, while in-service teachers were coded with the code T. Another measure taken to ensure the confidentiality of research data was keeping the university where the study was conducted undisclosed. This way, efforts were made to prevent the disclosure of both the participants' and university faculty members' identities.

3.9 Assumptions of the Study

This study was completed under some assumptions. Firstly, it was assumed that the multiple-choice questions and options in the assessment literacy inventory and interview questions were adequate, clear, and relevant to accomplish the research objectives. Secondly, it was assumed that the teacher education and assessment-related experiences of pre-service and in-service teachers who graduated

from the same university were similar. Thirdly, it was assumed that the participants answered the questions in the assessment literacy questionnaire and the questions asked in the interviews honestly and by reading in a way that made the most sense to them.

Fourth, it was assumed that the environments in which the participants attended the interviews with the online meeting tool were quiet, non-distracting, and appropriate. Finally, it was assumed that the researcher approached the participants during the data collection and administration phases without any pre-existing biases regarding their knowledge and experiences.

CHAPTER 4

FINDINGS

One of the aims of this study is to determine the assessment literacy levels of pre-service and in-service middle school mathematics teachers. The other aim is to get the participants' opinions about the deficiencies of the undergraduate education they received to the assessment literacy and their suggestions for improving the quality of assessment education in teacher education programs. This chapter provides descriptive statistics regarding the assessment literacy inventory, as well as qualitative analysis addressing deficiencies in the teacher education program in terms of developing assessment literacy. It also includes suggestions for enhancing assessment literacy within the context of a teacher education program based on interviews.

4.1 Descriptive Statistics

Assessment literacy inventory was applied to 34 pre-service and 41 in-service teachers before starting the interviews in the qualitative part of the study. This section gives descriptive statistics about pre-service and in-service teachers' assessment literacy levels. The total score of the assessment literacy inventory for pre-service teachers is 15.94 (SD = 2.56), while the total score for in-service teachers is 15.73 (SD = 3.03). The mean and standard deviation for assessment literacy levels of pre-service and in-service teachers are given in Table 4.1.

Table 4.1: Descriptive Statistics Results of Pre-service and In-service Teachers' Assessment Literacy Levels

Standards	**Pre-service Teachers (N = 34)**		**In-service Teachers (N = 41)**	
	Mean	**SD**	**Mean**	**SD**
ST1: Teachers need to possess the competence to select proper assessment techniques, tools, and methods for teaching decisions.	3.24	1.13	3.32	1.06
ST2: Teachers need to possess the competence to generate proper assessment techniques, tools, and methods for teaching decisions.	2.18	1.09	2.58	1.22
ST3: Teachers need to possess the competence to administer, score, and interpret the assessment results.	3.35	1.01	2.88	1
ST4: Teachers need to possess the competence to use assessment results for making decisions related to students, instructional planning, curriculum development, and school improvement.	2.21	0.8	2.15	0.88
ST5: Teachers need to possess the competence to create accurate techniques for grading students based on their assessments.	1.85	1.01	1.78	0.88
ST6: Teachers need to possess the competence to explain assessment findings to students, colleagues, and parents.	3.12	0.98	3.02	1.15
Total Score	15.94	2.56	15.73	3.03

Table 4.1 shows that the pre-service teachers got the highest score on ST3 (M=3.35, SD= 1.01) and the lowest score on ST5 (M=1.85, SD= 1.01). Similarly, table 4.1 also shows that in-service teachers obtained the highest score on ST1 (M=3.32, SD= 1.06) and the lowest score on ST5 (M=1.85, SD= 0.88).

When the scores of pre-service and in-service teachers were compared according to their proficiency areas, it was found that in-service teachers got higher scores in ST1 and ST2. In ST1, pre-service teachers (M=3.24, SD= 1.13) scored lower than in-service teachers (M=3.32, SD= 1.06). In ST2, pre-service teachers (M=2.18, SD= 1.09) scored lower than in-service teachers (M=2.58, SD= 1.22). On the other hand, pre-service teachers scored higher than in-service teachers in four standards, namely ST3, ST4, ST5 and ST6. In ST3, pre-service teachers (M=3.35, SD= 1.01) scored higher than in-service teachers (M=2.88, SD= 1.00). In ST4, pre-service teachers (M=2.21, SD= 0.80) scored higher than in-service teachers (M=2.15, SD= 0.88). In ST5, pre-service teachers (M=1.85, SD= 1.01) scored higher than in-service teachers (M=1.78, SD= 0.88). Lastly, pre-service teachers (M=3.12, SD= 0.98) scored higher than in-service teachers (M=3.02, SD= 1.15) in ST6.

4.2 Inferential Statistics

The independent t-test relies on three assumptions: the scores must be independent, exhibit normal distribution, and homogeneity of variance (Gravetter & Wallnau, 2014). Regarding the initial assumption, the scores were considered independent due to the data being gathered from two distinct groups, namely pre-service and in-service middle school mathematics teachers. The second assumption, which is normality, was tested using the Shapiro-Wilk Test, a method employed for small sample sizes with less than 50 samples, along with histogram analysis, skewness, and kurtosis. Skewness and kurtosis values are provided in Table 4.1 below. Since both pre-service and in-service middle school mathematics teachers' scores have skewness and kurtosis values within the range of 2 to -2, it can be assumed that these scores follow a normal distribution.

Table 4.2: Results of Skewness and Kurtosis Values

Group	N	Skewness	Kurtosis
Pre-service	34	-0.058	-0.276
In-service	41	-0.375	-0.002

Furthermore, the statistics for the Shapiro-Wilk Test is provided in Table 4.3. As seen in Table 4.3, the sigma value for both groups is greater than p = 0.05 according to the Shapiro-Wilk test statistics, confirming the normality assumptions for both pre-service and in-service teachers scores. Additionally, the histograms of these scores are included in Appendix D.

Table 4.3: Tests of Normality by The Shapiro-Wilk Test Statistics

Groups	Statistic	df	Sigma Value
Pre-service	0.973	34	0.555
In-service	0.966	41	0.248

For the third assumption, the Levene Test for Equality of Variances was employed to assess the homogeneity of variances. The Levene's test indicated that the variances in the assessment literacy inventory scores were equal for both pre-service and in-service middle school mathematics teachers, $F(1,75) = .956$, $p = .331$.

A null hypothesis was formulated to determine whether there is a significant difference between the mean assessment literacy scores of pre-service and in-service middle school mathematics teachers. This null hypothesis states, "There is no significant difference in the mean of assessment literacy inventory scores for pre-service and in-service teachers."

An independent-samples t-test was conducted to compare the assessment literacy scores between pre-service and in-service teachers. According to results, there was not a significant difference in the mean assessment literacy inventory scores between pre-service (M= 15.94, SD= 2.56) and in-service teachers (M= 15.73, SD= 3.03) conditions; $t(75)$= 0.319, p= 0.75. These results could imply that both pre-service

and in-service teachers might have similar gaps or limitations in their understanding of assessment concepts, techniques and practices.

4.3 Participants' Views About the Shortcomings of the Teacher Education Program Regarding Assessment Literacy

The study participants, including pre-service and in-service teachers, were questioned about the shortcomings of the teacher education program in enhancing assessment literacy. Specific areas where the program lacked were discussed for each of the six competency areas (standards) covered in the assessment literacy inventory, which was used as the quantitative data collection tool. Consequently, the areas that need improvement in undergraduate assessment education were analyzed and presented separately for each competency area. Additionally, certain shortcomings that were applicable to nearly all sub-dimensions were consolidated and labeled as general deficiencies. Deficiencies related to all other competency areas, including general deficiencies, were categorized as theoretical knowledge and practice/experience.

4.3.1 General Deficiencies

The general deficiencies of the teacher education program regarding assessment literacy development were divided into two categories: (i) theoretical knowledge and (ii) practice/experience. Table 4.4 shows these categories and sub-categories.

Table 4.4: General Deficiencies of the Teacher Education Program Regarding Assessment Literacy

Main Category	**Sub-category**	**Frequency**
Theoretical knowledge	The superficiality of the content of assessment related courses	9
	Gaps between theory and practice	8
	Inadequate number of elective and compulsory courses related to assessment	4
	Inability to grasp assessment concepts due to taking the assessment course too early in the program	4
Practice/ experience	Insufficient opportunities to apply theoretical knowledge of assessment in practice	10
	Putting the assessment into the background and not giving enough emphasis on it in teaching experience	6
	Insufficiency to receive feedback in the assessment parts of self-developed tasks	4
	Lack of teaching experience in highly diverse schools	4
	The internship experience falls short in terms of observing assessment process using by mentor teacher	4
	Inadequate assessment experience due to the assessment course not being student-centered	2
	Having delayed assessment practices despite learning the theoretical knowledge earlier	2

4.3.1.1 General Deficiencies in Theoretical Knowledge

Almost all participants mentioned the superficiality of the content of the assessment-related courses. They pointed out that assessment encompasses numerous sub-topics, but the program falls short in covering and teaching these topics. For instance, T2 expressed their thoughts as *"Because I remember that we didn't learn that much detail about assessment in the assessment course itself."* P3 supported what T2 said: *"I remember that assessment was covered very superficially in the classes. I think there was only reading and very little discussion."*

Many participants, especially in-service teachers, have mentioned the gap between theory and practice as a major deficiency of the program. They stated that what they learned in their undergraduate education about classroom environments, student levels, school culture, or student needs was based on ideal conditions, but real-life situations were different from what they learned in theory. Therefore, they mentioned facing assessment challenges when trying to apply their theoretical knowledge in real classrooms. For example, T1 explained the difficulty she experienced after becoming a teacher as follows:

You know when they said "Practice and theory are very different." It's really true in that regard. Because theory only gives us advice like "do it this way" or "do it that way." It doesn't provide anything else. But unfortunately, the reality is not like that

T4, similar to T1, expressed how the gap between theory and practice affected their teaching as follows:

We used to think that students had different levels, like low-level students, intermediate-level students, and good-level students. But I'm experiencing a situation based on where I live: there is no such thing as a good level. There are only intermediate and low levels. However, during our undergraduate education, we did not learn considering that we would have students at such a low level. Because I wasn't prepared for this, I am having a hard time assessing them.

Some participants pointed out the inadequacy of the teacher education program in terms of offering a limited number of elective and compulsory courses on

measurement and assessment. P1, who did not have the opportunity to take an adequate number of measurement and assessment courses during their undergraduate education, expressed their difficulty as follows: *"I took only one course during my undergraduate education, and I have actually forgotten a lot. If I had taken a few more courses, I would have felt more qualified."*

Another shortcoming mentioned by the participants was taking assessment lesson too early and not being able to grasp and elaborate on the assessment concepts. They highlighted that at the time they took the assessment course, they did not have sufficient knowledge about mathematics education, which hindered their ability to grasp the material deeply. Furthermore, they stated that taking the assessment course early resulted in forgetting its content. T2 articulated this shortcoming as follows:

We took the assessment course in the second grade. You don't know about lesson plans in the second grade, you don't know how to design a lesson, or you don't know how mathematics is taught. We haven't even reached that stage. But there is a course being taught on assessment techniques. I mean, I'm thinking I wasn't very conscious.

4.3.1.2 General Deficiencies in Practice/experience

All of the participants remarked on the insufficient opportunities to apply theoretical knowledge of assessment in practice. The participants mentioned that they were unable to translate their theoretical knowledge into practice in areas such as selecting/developing appropriate assessment methods, scoring and interpreting assessment results, planning instruction based on assessment results, and communicating with parents, students, or other educators regarding assessment results. They expressed that without being able to put their theoretical knowledge into practice, they couldn't test their own learning and they felt that the theoretical knowledge they acquired remained abstract. They also noted that even their teaching experiences involved more observation than actual teaching practices. P5 shared his thoughts about feeling inadequate in assessment due to not having enough opportunities for practical application in assessment:

We don't have any experience in applying what we have learned in real life. In other

words, we don't have experience. I think this causes a bit of a problem: it leaves the subject hanging. Also, I don't feel competent enough in assessment since I haven't had such an experience before. Therefore, I need to gain more experience from my own perspective. I also feel that there are forgotten knowledge and processes that I have experienced during the time I haven't practiced. I feel this as I talk about it.

T4, on the other hand, mentioned the limited opportunity to apply their knowledge about assessment during the teaching experience and its consequences as follows:

Well, during the internship course, we didn't have much opportunity for practice. I mean, throughout the four months, I only taught a class once. I know most of us (teacher candidates) are in the same situation. The internship course was mainly focused on observation, there wasn't much practical application. So, when there is not much opportunity for practice, it doesn't have much impact. Observation is important of course, but there was a lack of practical application.

Putting assessment into the background and not giving it enough emphasis is another deficiency mentioned by more than half of the participants. They noted that measurement and assessment are often overlooked in tasks such as lesson planning or teaching experiences, resulting in a lack of importance placed on assessment. T2 provided an example based on her experiences:

I think things related to measurement and assessment are falling behind a bit. We focused more on how to deliver lessons, how to teach, and how to manage the classroom. We did put assessment into the background. For example, we often spent more time designing activities for assignments and neglected the assessment aspect.

P5 is another participant who referred to this deficiency:

The internship is a few weeks long, and we have a 40-minute teaching experience. During this time, we tended to think that it is enough to implement our lesson plans and activities. Moreover, for someone who has never done an assessment before, it can be challenging to incorporate assessment into lesson plans. As a result, we often overlooked the assessment part and focused on creating better activities or engaging in more enjoyable activities.

Another deficiency experienced by the participants in the teacher education program is insufficiency to receive feedback in the assessment parts of self-developed tasks. Almost half of the participants mentioned that they received feedback from university faculty members for some of the tasks they prepared, but they said that this feedback lacked in terms of assessment. As a result, participants stated that they could not gain information about their strengths and weaknesses in assessment. This has led to a deficiency in identifying areas they need to improve and making necessary adjustments in their assessment strategies. For example, P3 stated: *"I don't have a clear idea about whether the questions I have prepared will be beneficial or useful for students because I didn't receive specific feedback on the questions I created."*

Participants also indicated that the lack of diversity in the schools they visited during their teaching experience was another deficiency of the teacher education program. They noted that the schools they were placed in had similar cultures and student profiles, which limited their exposure to different classroom needs and experiences. T3 explained this deficiency:

During the internship, we all worked in a specific region, meaning we were in the same schools. We didn't get to see many different places. Therefore, to be honest, we didn't get to see where assessment methods (in what types of classrooms) worked or didn't work much. This was also a deficiency we experienced.

The internship experience falls short in terms of observing the assessment process used by mentor teachers is another deficiency that participants specified. The participants mentioned that they did not observe or discuss with mentor teachers which assessment methods they used, how they scored and interpreted assessment results, how they incorporated assessment results into their lesson planning, and how they conducted valid grading. For instance, T3 elaborated on his experience:

During the internship, we observed [a classroom], but I don't think it contributed much in terms of assessment. It mainly allowed us to observe the school environment and surroundings. Of course, it might have offered something in terms of teaching and mathematics education, but I didn't see many examples of assessment in the teachers I observed. Therefore, I can't say it made much of a contribution in that aspect.

T5 and P3 pointed out the deficiency of the inadequate assessment experience due to the assessment course not being student-centered. They mentioned that the lessons generally progressed through slides and had a predominantly teacher-centered approach. In other words, they highlighted the lack of teacher-student interaction in the assessment course and stated that teacher candidates had low engagement in the class. For instance, P3 expressed that:

In the assessment course, we were progressing through slides. Generally, the instructor himself/herself was solving the questions. Actually, I think there could have been an opportunity for teacher candidates to solve them as well. It didn't quite resonate with us when the teacher was the only one speaking.

Lastly, two participants mentioned that they had opportunities for assessment practices but were late in the program despite learning the theoretical knowledge on assessment earlier. For instance, P2 explained this deficiency as follows:

We were very late in applying what we had learned. When we went for the internship, we were already in the fourth year, so I think it was even too late to assess our competence in this regard. There might be such a deficiency.

4.3.2 Deficiencies in the Selecting and Generating Appropriate Assessment Methods

The deficiencies declared by the participants in the 1st and 2nd competency areas, specifically in selecting and developing suitable assessment methods, were classified into two categories: (i) theoretical knowledge and (ii) practical experience. Table 4.5 presents these categories, sub-categories, and the shortcomings of the teacher education program along with their frequencies.

Table 4.5: Deficiencies of the Teacher Education Program in Selecting and Generating Proper Assessment Methods

Main Category	**Sub-category**	**Frequency**
Theoretical knowledge	Lack of knowledge in choosing/developing an assessment that can meet the diverse classroom needs	7
	Inadequate knowledge on assessment preparation procedure and process	2
	Lack of knowledge in developing an assessment that attends to the educational and contextual difficulties	2
	Inability to get prepared to overcome language difficulties in assessment	1
Practice/ experience	Limited exposure to a variety of assessment methods and tools/tasks on assessment	2

4.3.2.1 Theoretical Deficiencies

Most participants reported that the teacher education program fell short of providing the necessary theoretical knowledge in choosing/developing an assessment that can meet diverse classroom needs. The participants mentioned that their undergraduate education was insufficient in determining an assessment method that can meet the needs of the entire class, which usually includes students with different academic levels (good-average-poor) or gifted and disabled students. For instance, P4 expressed his thoughts as follows: *"How should I approach assessment in a class of 20 or 30 students? I don't know this. Because their needs can be different. Their ways of expressing themselves can be different."*

P2 expressed that despite the existence of inclusive students within schools, the teacher education pertaining to evaluating these students is deficient, as stated in the following words:

Especially in our internship classroom, there were a lot of inclusive education students. Previously, this number was not as common, but now almost every class has at least two inclusive education students. As a teacher, you are delivering lessons and expecting them to comprehend at the same level. You also need to prepare an exam suitable for such students. Of course, you can receive support from the guidance counselor, but in terms of assessment, for example, as a teacher, you want to observe what these students understand. However, we don't know what level the student is at and what they can do. How should I provide an assessment for this? I don't know. I definitely think this is a deficiency.

Inadequate knowledge on assessment preparation procedures and processes is another deficiency of the teacher education program that participants noted. Participants mentioned that there is a lack of knowledge in terms of what to consider when preparing an assessment, such as how the distribution of questions should be based on topics or the distribution of difficulty levels for different types of questions. T2 expressed her thoughts as follows:

In undergraduate education, I believe there was a deficiency in addressing what we should pay attention to when preparing an exam. For instance, determining the appropriate number of questions for a specific topic or learning outcome, or how the question distribution should be managed. I still don't think I know them very well. I think there were shortcomings in this regard.

Two of the participants (P3 and P4) cited the shortcomings of the teacher education program in presenting them with the necessary knowledge to develop an assessment tool that attends to educational and contextual difficulties. The participants mentioned that they were facing various challenges such as a dense curriculum, numerous responsibilities for teachers, and classroom management issues in crowded classrooms, and they do not have an understanding of how to select/develop an appropriate assessment method.

T4, who was teaching in one of the rural schools, stated that he felt unready to handle language-related challenges in assessments in the teacher education program. He mentioned that the language spoken by the students in the school where he was working was significantly different from his own, and therefore, he struggled

in preparing assessments for them. He specifically mentioned having difficulty in creating valid and reliable assessments:

Here, there is a problem, a language issue. I don't remember if I learned about this language issue during my undergraduate education. Even if I did, I don't recall it being emphasized extensively. To overcome this problem, I try to choose my words carefully to ensure clarity in assessments. And to ensure validity. Sometimes, I provide additional explanations, but it still doesn't work. For example, right now we are talking and understanding each other, but unfortunately, it is not the same in school.

4.3.2.2 Practice/experience Deficiencies

Two of the participants, T2 and P2, expressed that limited exposure to a variety of assessment methods and tools/tasks was a deficiency of their undergraduate education. The participants mentioned that teacher candidates tended to use similar methods in tasks that require the preparation of assessment methods, and therefore, they were unable to observe different methods. In other words, they stated that they were unable to expand their perspectives and ideas regarding assessment methods. P2 described the shortcoming she encountered in this regard:

Most of us (teacher candidates) were using similar assessment methods. For example, if we were 20 people, then 15 people did the same thing. Maybe, if any, five of us were following a different approach. As a result, the assessment methods used were always uniform.

4.3.3 Deficiencies in the Administering, Scoring, and Interpreting the Assessment Results

The results related to competency area 3, which is administering, scoring, and interpreting assessment results, were reported under two main categories: (i) theoretical knowledge and (ii) practice/experience. Table 4.6 shows these categories and sub-categories that cover the deficiencies of the teacher education program and

the number of participants who reported them.

Table 4.6: Deficiencies of the Teacher Education Program in Administering, Scoring, and Interpreting the Assessment Results

Main Category	**Sub-category**	**Frequency**
Theoretical knowledge	The irrelevance of the tasks used in the statistics course for an assessment in an educational context	3
	Lack of knowledge on identifying students' deficiencies based on assessment results	3
	Lack of knowledge on scoring unexpected answers from students in assessments	2
	Insufficient knowledge on interpreting assessment results for a larger student population	2
	The rubrics taught in the lessons were too superficial and did not provide an in-depth understanding	1
Practice/ experience	Lack of practice on interpreting assessment results	5
	The insufficient practice of newly acquired statistical concepts applied	1
	Insufficient experience in applying a rubric to assess actual student work (or students' actual work)	1

4.3.3.1 Theoretical Deficiencies

One of the deficiencies of the teacher education program that the participants highlighted was the irrelevance of the tasks used in the statistics course for an assessment in an educational context. They have pointed out that the examples, solved problems, or exercises given in the statistics course were not relevant in an educational context, specifically in terms of assessment. For this reason, they said that it did not help them to establish a connection between statistics and assessment. Thus, they have emphasized the inadequacy of the statistics course in managing and interpreting assessment results. For instance, T2 shared her thoughts considering her

experiences:

I had reached a point of rebellion, questioning why we were taking statistics. Because I didn't know how the statistical knowledge, I learned in class would be useful to me. For example, how I would interpret students' results in this way, or how I would evaluate the outcome when conducting a large-scale study or assessment in the future? Because the context was very different, as I mentioned. The contexts used by the instructor were mostly related to business topics. It had nothing to do with education. Perhaps, due to this gap and lack of connection, I believe it fell short.

Another shortcoming mentioned by the participants is the lack of knowledge on identifying students' deficiencies based on assessment results. Three participants emphasized that the contribution of undergraduate education is limited in terms of providing them with the necessary knowledge to identify what students have learned well or did not learn, as well as their misconceptions, especially in formative assessments. They highlighted the need for a greater emphasis on these aspects in the undergraduate teacher education program.

Two participants, T2 and P4, stated that the teacher education program fell short of providing knowledge on scoring unexpected answers from students in assessments. Participants asserted that they did not know what to do about and how to score and interpret a solution that they had never encountered before. As P4 put it :

What if an unexpected solution comes up? What will I do then? Let's say I have set criteria like if they do it this way, full points or if that is missing, one point deducted. But what if a student provides a completely different answer, a different solution? I do not know how to evaluate it.

Two participants referred to the insufficient training on interpreting assessment results for a larger student population as a deficiency of the teacher education program. They stated that they have received insufficient knowledge on how to compare the achievements and learning processes of students in overcrowded classrooms or across different classes. T2 shared her point of view:

I now know how to compare the success between them in cases where I need to provide more comprehensive assessments, such as in larger student groups or when

I have three classes. However, I gained this knowledge through teaching, and my undergraduate education did not contribute much to this regard. If I wasn't a teacher, the knowledge I gained in my undergraduate education would not have been enough to do this. I believe there were deficiencies in these areas.

Another issue that one of the participants pointed as a deficiency was related to the rubrics. The participant thought that that the fundamental concepts of constructing a rubric taught in the assessment course were superficial and did not facilitate a comprehensive comprehension. She stated that they often struggled with the rubrics they prepared for homework, and they were unsure if they were doing it correctly. P2 expressed her thoughts as follows:

The rubrics we saw in undergraduate education were very superficial. That's why we often struggled when preparing our own rubrics. For example, I would have doubts like "Did I evaluate it correctly?"

4.3.3.2 Practice/experience Deficiencies

One of the participants, P3, remarked on the insufficient practice of newly acquired statistical concepts applied in the teacher education program. The participant mentioned that in the statistics course, the instructor always solved the problems themselves, and the teacher candidates were not given such an opportunity. Therefore, they talked about their inability to improve their knowledge of statistics. P3 said:

In the class, we were only progressing through slides, for example. Our teacher solved the questions themselves. For instance, we learned things like calculating the mean or standard deviation. But we were not given an opportunity to solve questions ourselves and practice on them.

In terms of practices related to teaching, P4 asserted that undergraduate education did not offer them sufficient experience in applying a rubric to assess actual student work (or students' actual work). She noted that even though they prepared rubrics, these rubrics remained very theoretical because they couldn't apply them in a real classroom with actual student work. P4 elaborated on this deficiency by sharing this experience:

We prepared assessments and rubrics in our lesson plans. However, we didn't actually collect the exams or assessments we conducted in class and evaluate them according to those rubrics.

4.3.4 Deficiencies in the Using Assessment Results for Making Educational Decisions

The results related to competency area 4, which is making decisions about students, planning the teaching, developing the curriculum, and using the evaluation results for school improvement were reported under two main categories: (i) theoretical knowledge and (ii) practice/experience. Table 4.7 shows these categories and sub-categories that cover the deficiencies of the teacher education program in this competency and the number of participants who reported them.

Table 4.7: Deficiencies of the Teacher Education Program in Using Assessment Results for Making Educational Decisions

Main Category	**Sub-category**	**Frequency**
Theoretical Knowledge	Lack of knowledge about taking individual differences into account when making decisions based on assessment results	5
	Lack of addressing what can be done to improve teaching after assessment is applied	3
Practice/ experience	Insufficient practice in planning the subsequent lesson according to the results of the assessment	7

4.3.4.1 Theoretical Deficiencies

Half of the participants remarked on the shortcomings of the teacher education program in providing knowledge about taking individual differences into account when making decisions based on assessment results. Participants emphasized that

all students in a classroom have different personalities and levels, resulting in diverse needs. They mentioned that while the needs in a classroom environment vary for each student, they lacked knowledge on how to make the most appropriate decisions for all students after conducting assessments. T3 described the difficulty he was having in this respect:

In this regard (planning lessons using assessment results), I don't remember anything, and it must be discussed during undergraduate studies. For example, I don't remember reading any academic papers or having discussions on this topic in undergraduate education. We always got stuck in this aspect. This is the greatest truth of life: students are different. We always say they are different, but how will we interpret their differences? After interpretation, how will we actively implement it, take precautions, or provide support? We never discussed or demonstrated various methods for that.

Another issue that participants experienced was an absence of in terms of theoretical knowledge in their undergraduate education is the lack of addressing what can be done to improve teaching after the assessment is applied. The participants mentioned that their undergraduate education was incomplete, particularly in terms of addressing what actions can be taken when the desired success in the classroom cannot be achieved. Participants mentioned that there is a curriculum that teachers need to be trained in, but they do not know what can be done within a limited time. For instance, P4 cited this shortcoming as:

Let's say the students didn't understand. According to the assessment results, they didn't grasp a certain topic. Now, when I think about it, okay, I can go back and explain that topic again. But, you know, how I will catch up on the topic afterward is also important. Of course, it is crucial that they understand the topic first. I definitely need to go back to that. But I need to further develop myself on how to do it. Should I consider additional lessons or is it logical to explain two topics at the same time? I don't know what I would do if most of the class didn't understand, and I feel overwhelmed about not being able to catch up. Frankly, I feel inadequate in this regard.

4.3.4.2 Practice/experience Deficiencies

The majority of the participants view the teacher education program's failure to provide adequate opportunities for teacher candidates to incorporate assessment results into their planning for the next lesson. They mentioned that both in their internship experiences and lesson plans, they focused on preparing for the one-time lesson and didn't consider the next lesson. They also noted that they did not prepare a lesson plan using the assessment results. As a result, they stated that they couldn't gain practical experience in planning the next lesson based on assessment results. For example, T1 exemplified that with his experiences:

This is a deficiency in undergraduate education. We weren't preparing lesson plans based on any assessment results. We were given a topic, and we were only focusing on that topic.

Additionally, P3 expressed shortcomings in planning the lesson using the assessment results with the following words:

During the internship course, we prepared many lesson plans, but each of us was only responsible for one lesson. The next lesson was taught by another prospective teacher or an actual teacher. Therefore, we didn't have the opportunity to identify students' shortcomings, address them, and create plans based on that. We simply taught what was expected of us in that specific lesson. So, even if there were gaps in students' understanding, it wasn't our responsibility to address them in the next lesson. The responsibility shifted to whoever was teaching the next lesson. That's why we didn't have the chance to use assessment results for lesson planning since we only taught for one hour.

4.3.5 Deficiencies in the Creating Accurate Techniques for Grading

The participants have identified the shortcomings of their undergraduate education in developing valid pupil grading procedures. There is only one category for this deficiency: theoretical knowledge. Table 4.8 shows this category, sub-categories and frequencies.

Table 4.8: Deficiencies of the Teacher Education Program in Creating Accurate Techniques for Grading

Main Category	Sub-category	Frequency
Theoretical knowledge	Theoretically bypassing the valid grading process	7
	The inadequacy of the teacher education program in presenting the realities of the Turkish education and evaluation system for a valid grading	2
	Lack of knowledge about how to handle classroom performance grade in the valid grading process	1

4.3.5.1 Theoretical Deficiencies

Many participants have mentioned that they theoretically bypassed the valid grading process during their undergraduate education, and as a result, they were not able to follow a procedure to give their grading. For instance, they lack knowledge about which assessments to consider when assigning end-of-term grades or how much weight each assessment should carry. They expressed a lack of understanding regarding the process that should be followed in determining the impact of different assessments. T2 explained the deficiencies of undergraduate education about the current grading procedure as follows:

I don't remember discussing how to give valid grades to student performance or how to make final decisions. We mainly talked about not focusing solely on the outcomes but also evaluating the process. We discussed these general aspects, but I don't recall any specific discussions about how to actually assign or determine grades.

The inadequacy of the teacher education program in presenting the realities of the Turkish education and evaluation system for a valid grading was another deficiency stated by some of the in-service teachers. They mentioned that despite their attempts to give valid grades to their students based on the knowledge they acquired during

their undergraduate education, they were unable to do so according to the Turkish education and evaluation system. They stated that the teacher education program did not provide them with information in this regard, and as a result, they faced difficulties in valid grading. T4 described the difficulty he was having in this respect:

In undergraduate education, we had various types of grades to give to students. However, in the national education system, if a student's written exam score is 80, you cannot give them a performance grade lower than 80. I'm not sure if this is a deficiency in undergraduate education, but this aspect was not considered. Let's say the student is performing very poorly in class, unable to do anything, doesn't know anything, but let's say they cheated on the written exam and got an 80. We cannot give them a lower grade. We cannot give them the grade they truly deserve. In this regard, undergraduate education should have been aware of certain aspects in the national education system.

Another way of grading that teachers in Turkey are obligated to give at the end of the semester is classroom performance grades which encompass various factors ranging from the student's overall performance to their behavior in the classroom. One of the participants, P2, highlighted that the guidance provided by undergraduate education on grading students' general classroom performance in a valid way was limited.

4.3.6 Deficiencies in the Explaining Assessment Findings to Students, Colleagues, and Parents

Participants identified several deficiencies in their teacher education program to equip them with the necessary knowledge and skills to communicate assessment results to students, parents, and other educators. These deficiencies were grouped into two categories: (i) theoretical knowledge and (ii) practice/experience. Table 4.9 illustrates these categories and sub-categories, as well as the frequency of areas that need improvement in undergraduate education.

Table 4.9: Deficiencies of the Teacher Education Program in Explaining Assessment Findings to Students, Colleagues, and Parents

Main Category	**Sub-category**	**Frequency**
Theoretical knowledge	Inadequate preparation in communicating assessment results to parents	8
	Inadequate preparation in communicating assessment results to students	6
	Inadequate preparation in communicating assessment results to colleagues	4
	Inadequate preparation in formal reporting to defend assessment results	2
Practice/ experience	Communication difficulties due to the perceptions of the students in the internship school toward pre-service teachers	3

4.3.6.1 Theoretical Deficiencies

Most of the participants perceived the teacher education program as lacking adequate preparation in communicating assessment results to parents. Although they acquired some information regarding parent communication during their undergraduate education, they pointed out that it was not about assessment results. T1 expressed the deficiency in undergraduate education regarding this matter based on his current teaching experience as follows:

In my opinion, undergraduate education has contributed nothing in this regard. Throughout my undergraduate education, I never received any training on how to communicate with parents or how to discuss assessment results with them. I always feel the absence of this knowledge because I never developed such a habit. This, in my view, is one of the biggest deficiencies of my undergraduate experience.

P1 expressed that they saw the subject of communication with parents in the classroom management course, but they did not associate it with assessment:

We only talked about communication in the classroom management class, teacher. However, in the classroom management course, there was no specific subject that directly focused on assessment. In other words, we learned how to communicate with parents and students, but not specifically how to communicate about evaluation results. For example, we discussed scenarios where we share written results and the parent reacts, or when a student is unhappy and cries. We discussed how to handle such situations. But to the best of my recollection, there was no specific topic that directly addressed the communication of evaluation results.

Another deficiency mentioned by the participants was inadequate preparation in communicating assessment results to students. More than half of the participants pointed out that, similar to communication with parents, what they learned about communicating with students was not assessment-based, and that their undergraduate education did not provide much guidance on how to convey assessment results and what steps to take in that regard. Two of the participants have also mentioned that undergraduate education fell short in preparing them for communicating assessment results to their colleagues. For instance, P4 said that: *"My undergraduate education did not provide me with any benefits regarding how I would communicate and collaborate with other teachers if I were to become a teacher."*

Two participants (T2 and P1) argued that their preparation in formal reporting to defend assessment results was inadequate. They mentioned that they are aware of the importance of constantly communicating with students, parents, and colleagues about assessment results for the student's development. However, they expressed that they don't know how to formalize this process and document it effectively. For instance, T2 highlighted her difficulty regarding the formal reporting of assessment results:

Yes, we evaluate students' learning process, but being able to report on the student's overall situation continuously and formally, identifying their specific difficulties or fluctuations, is something I struggle with. Honestly, I feel I have a deficiency in this aspect. Our undergraduate education might have fallen short in terms of documenting and reporting assessment results.

4.3.6.2 Practice/experience Deficiencies

The shortcomings identified by the participants were communication difficulties due to the perceptions of the students in the internship school toward pre-service teachers. The participants mentioned that, during their teaching experience, they had the opportunity to interact and communicate with students. However, they pointed out that the students did not perceive them as real teachers and, as a result, did not take them seriously. Therefore, they stated that the teaching experience did not contribute significantly to communication, including assessment, with students. P5 illustrated this with his own experiences:

In the internship class, students perceive teacher candidates as older brothers/sister. The sooner you overcome this perception, the better, of course. However, you see, elementary school students especially have only one teacher. They embrace them and they are the authority figures for them. When you go there as an intern, they like you too, but you can't go beyond being an older sibling to them.

According to the pre-service and in-service teachers, in the internship schools, they were not seen as a teacher with all teaching responsibilities. This may show the inadequacy of the cooperation between the university and other stakeholders (mentor teacher, school administration, and students) to ensure healthy communication of teacher candidates with students within the scope of the internship.

4.4 Suggestions for Enhancing Assessment Literacy in a Teacher Education Program

Pre-service and in-service teachers were asked for their suggestions on how the teacher education program could be more effective in improving their assessment literacy. The suggestions were analyzed and categorized into three main themes based on the framework of Juanjuan and Mohd Yusoff (2022) for an effective professional development program to improve the assessment literacy of teachers. The first category focused on the "what" aspect of the suggestions that were related to the content of the teacher education program on assessment. The second category

focused on the "how" aspect of the suggestions that were related to the ways of developing pre-service teachers' assessment literacy. The third category focused on the "why" aspect of the suggestions that indicated the teacher education program's underlying orientation in developing teachers' assessment literacy. In the next section, these suggestions are presented within these three main themes.

4.4.1 What: Content of a Program to Improve Assessment Literacy

Pre-service and in-service teachers' suggestions for improving assessment literacy in teacher education were categorized based on the success factors outlined in Juanjuan and Mohd Yusoff (2022). In this respect, the suggestions of the participants were examined and grouped into three main categories as follows: (i) The practical application of assessment knowledge and skills, (ii) The structure and sequencing of assessment-related courses in the teacher education program, and (iii) Enriching the content of assessment courses according to assessment difficulties and daily assessment practices. Table 4.10 shows pre-service and in-service teachers' suggestions regarding the content of the teacher education program that were listed under three main categories and 13 sub-categories.

4.4.1.1 Practical Application of Assessment Knowledge and Skills

The participants offered many suggestions to make the assessment education given in the teacher education program more effective in terms of content. Most participants stated that teacher candidates should gain and develop assessment literacy by focusing on more assessment practices. Hence, several of their suggestions were related to creating opportunities for the application of theoretical knowledge in actual assessment contexts.

Most participants emphasized that the courses should include various assessment-related assignments that would give opportunities for them to apply the theoretical knowledge they learned in undergraduate education. In this direction, they suggested increasing the scope and number of assignments related to assessment. For example, T5 said:

Table 4.10: Suggestions For Enhancing Assessment Literacy in the Content of Teacher Education Program

Main Category	Sub-category	Frequency	a
Practical application of assessment knowledge and skills	The courses should incorporate a diverse range of assessment-related assignments for pre-service teachers	9	
	The courses need to include actual or hypothetical student assessment results	5	
	The courses need to place greater emphasis on incorporating the assessment part into lesson planning	3	
	The program should require pre-service teachers to practice a wide variety of assessment methods	1	
The structure and sequencing of assessment related courses in the teacher education program	Increasing the number of compulsory and elective courses related to assessment	9	
	Taking the assessment course with greater awareness in the later stages of the undergraduate program	3	
	Configuring the program so that the connections between assessment-related concepts learned in different courses can be established	4	
Enriching the content of assessment courses in line with assessment challenges and daily assessment practices	Gaining more information about how the grading process should be	6	
	Gaining knowledge about how to plan instruction according to different assessment results	3	
	Customizing the theoretical knowledge provided in the courses according to the needs of the students	2	
	Gaining knowledge about how to interpret different solutions for the same assessment	1	
	Gaining knowledge on how to connect classroom management, lesson planning and assessment	1	
	Gaining information about school administration and its role in the assessment process	1	

What I mean is, it would be better if I could see my own mistakes in my assignments. And, as I mentioned, I probably saw them in the assignments I did back then (during my undergraduate studies). But why shouldn't there be a greater number of these assignments? I can say that. Because, for example, in addition to talking about how you can choose or develop your questions, there is also the option of actually doing it and seeing the results.

In addition to T5, T1 also mentioned the importance of assessment-related assignments:

I mentioned earlier that we had prepared questions for an exam (as an assignment), and I had created an analytic rubric for those questions. Our instructor evaluated that analytic rubric and provided feedback. That incident was very helpful to me. What would be my advice in this regard? I wish we had prepared not just one but five assignments like this. We would have learned much better.

Another suggestion voiced by participants was that courses should include actual or hypothetical student assessment results. They mentioned that incorporating actual or hypothetical student assessment results in courses would provide concrete examples and practical applications of the concepts they are learning. By doing so, they stated that teacher candidates could see how assessment results were obtained and interpreted, allowing them to understand better the subject matter and its relevance in real-world scenarios. For example, T2 voiced her opinion in these words:

What if we had actual or hypothetical exam results or cases and we made comments about assessment based on these results? Then, we could make comments on how to interpret these results or discuss topics such as designing activities for a group of students who obtained such results.

Participants also claimed that prospective teachers should be asked to include a detailed assessment section in their lessons. They argued that by integrating the assessment process into lesson planning, prospective teachers would learn assessment more effectively. Therefore, the participants' suggestion was that the courses need to place greater emphasis on incorporating the assessment part into lesson planning. P5 made the following recommendation on this issue:

In the internship course, prospective teachers may deliver a lesson in which the assessment methods are effectively integrated into the lesson plan they have created or will create. It is not just about creating a lesson plan, but creating a lesson plan that is suitable for assessment and using appropriate assessment methods to measure and evaluate students' knowledge within a 40-minute timeframe can be requested as one of the lesson plans.

T2 supported the views of P5 and stated that:

We often think that assessment occurs only at the end of a lesson, but it is actually a continuous process. Maybe more emphasis can be placed on assessment in lesson plans. It would be beneficial to include both different assessment methods and incorporate activities or anecdotes between lessons that facilitate assessment.

One of the participants (P2) suggested that the program should require preservice teachers to practice a wide variety of assessment methods. The participant mentioned that in tasks requiring assessment, most prospective teachers use the same methods, but she emphasized that utilizing different assessment methods would enhance their perspectives. She explained her ideas by expressing:

We can be open to innovations. For example, after a lesson, I don't have to distribute activity sheets and look at what the child has gained from them... What if, let's say, three people use a specific method to implement their assessment and conduct their test. The others would apply different assessment methods. We would then discuss the results we obtained. We would analyze which approach was more effective for the child in the classroom. It didn't usually happen this way. We all did similar things. It was like just changing the questions. Our actions were very parallel to each other.

4.4.1.2 The Structure and Sequencing of Assessment Related Courses in The Teacher Education Program

Both pre-service and in-service teachers made some suggestions about the structure and order of the courses related to assessment in the teacher education program based on their experiences in undergraduate education.

One of the most mentioned suggestions is increasing the number of compulsory and elective courses related to assessment. The participants mentioned that there are many important subtopics included in the assessment dimension. However, they pointed out that the courses offered in the teacher education program are insufficient in covering these topics comprehensively. Therefore, they made suggestions to increase the number of courses in order to learn assessment in a more detailed and effective manner. The responses from T3 and P3 serve as an illustration of what these participants advised:

T3: *In a topic with so many subtopics (which is assessment), I believe there should be at least 5 elective courses dedicated to each of these subtopics. In fact, a separate course could be opened for each individual topic. Because it is very difficult to explain all of them in a short time or in one lesson. If you don't break down these topics into smaller parts over the course of 4 years, it would be difficult to teach them effectively and apply them in practice.*

P3: *I have never seen an elective course specifically focused on assessment, so I think it was insufficient in that aspect. In my opinion, there could have been at least two elective courses on assessment. In other words, I believe the number of elective courses on assessment could be increased, as there are students who want to improve themselves in this area. For example, instead of taking an elective course from another department, I would have preferred to take an elective course from my own department where I could enhance my skills in assessment. I think if there were an elective course on assessment, I would definitely have taken it, but I have not come across one before. Therefore, the number of assessment courses could be increased.*

Another suggestion mentioned by pre-service and in-service teachers was to consciously take the assessment course in the later stages of the undergraduate program. The participants noted that they had taken the assessment course at the beginning of their undergraduate studies and during that time, they were unable to establish a strong foundation of knowledge and skills in their respective areas of study. Therefore, they recommended taking the assessment course at a later stage when they have a deeper awareness of their profession. T1, for instance, expressed his thought as follows:

It would have been better for me if I had taken the assessment course at the beginning or toward the end of the third year. Because I started feeling more like a teacher around the end of the third year. If I had taken the assessment course around the end of the third year, I would have had a better understanding of what measurement and assessment mean or how to effectively use assessment concepts.

Three participants suggested that the program should be configured in a way that would allow for establishing connections between assessment-related concepts learned in different courses. The participants mentioned that they learned about assessment-related knowledge and skills in some other courses besides the assessment course, but they did not make connections between the knowledge and skills acquired in different courses. In this regard, they recommended that the assessment dimension should be emphasized, and its connections should be established in other courses other than the assessment course. T2 referred to this as:

It could be about putting together those puzzle pieces that will enable meaningful learning. Because they all influence each other. In fact, all of these courses are linked to each other. For example, we had already learned the necessary information regarding lesson design. We learned how to prepare a lesson plan, how the flow should be, and so on. However, we didn't connect this knowledge to the assessment aspect. Maybe if these learned pieces of information are combined in a coherent way, it could be more beneficial.

One of the teacher candidates made a suggestion that the content of the statistics course should include context related to assessment. The participant mentioned that they couldn't establish a connection between the information they learned in the statistics course during their teacher education program and assessment, resulting in poor data analysis and interpretation. The participant mentioned that providing examples in the statistics course that are more focused on the context of assessment would enable teacher candidates to conduct higher quality assessments when they become teachers. T2 referred to this as:

If we had received education from a professor in the field of education, just like we did in statistics during our graduate studies, and had approached it in the context of education, perhaps our perspective would have been different. It would have been

more comprehensible in that way.

4.4.1.3 Enriching The Content of Assessment Courses in Line with Assessment Challenges and Daily Assessment Practices

As mentioned earlier, the participants highlighted the insufficiency of theoretical knowledge in some aspects of assessment literacy during their undergraduate education. To address these deficiencies, they provided recommendations for the content of assessment education offered in teacher education programs. These suggestions were associated with enriching the content of assessment courses in line with assessment challenges and daily assessment practices.

The participants' most common suggestion was related to gaining more information about how the grading process should be. Many participants expressed that undergraduate education has little contribution to giving valid grades to students, and they considered it a significant deficiency. They emphasized the need for a teacher education program to provide more comprehensive education on the grading process, how to assess students through an oral exam, and how to make decisions related to students' progress to the next level (grade). They suggested that the program should take into account the grading process and offer enhanced training in these areas. T2 pointed this out by asking for answers to the following questions:

I believe it is necessary to address the following grading topics: How is student performance evaluated? How is it scored? What criteria should be considered while scoring? What should be the percentage of in-class performance? Or what factors should be considered when assessing in-class performance?

T3's thoughts on the need for the undergraduate education to provide more information about the criteria to be considered during grading aligned with the recommendations of T2:

I don't remember if we covered grading scales for overall assessment at the end of the semester. For example, I don't remember we were discussing the impact of specific criteria on grading or talking about why certain factors have more influence. I can provide a concrete example, let's say the first exam carries a weight of 50%

in grading, while the second exam carries 30%. Why is it structured this way? Can these weights vary? If yes, what factors contribute to the variability? These topics were either not discussed or I don't remember. It would be beneficial to give more emphasis to this aspect.

Some of the participants highlighted that the program should provide knowledge on how to plan instruction based on different assessment results. They mentioned a lack of understanding regarding what steps to take and how to plan lessons, particularly when desired assessment results were not achieved. They emphasized the importance of undergraduate education offering more information on lesson planning in relation to assessment outcomes to address this knowledge gap. P5 commented on this issue as follows:

Let's say that as a result of the assessment we conducted, we observed that the class is not performing at the desired level. In such a situation, I believe we need more in-depth theoretical knowledge about what actions to take. Because otherwise, experiencing the scenario where the class fails becomes a bit more dependent on chance. Instead of leaving this matter to chance, I think it is necessary to discuss it theoretically.

Two of the participants (T4 and P4) mentioned that customizing the theoretical knowledge provided in the courses to meet the specific needs of the students should also be included in the program. The participants emphasized that the social and cultural structure of the schools they will work in may vary in real life, and there may be students with different needs within a classroom. They argued that undergraduate education should train prospective teachers more effectively in this regard, so they can be more competent in addressing such diverse situations. For example, T4, a teacher working in a village school, made the following suggestion:

We are unable to apply the same assessments we learned in undergraduate education to the children here. This is because we need to tailor the assessments to the students' levels. Additionally, language becomes an issue when the geography is different. The language used in the assessments needs to be simplified. For example, sometimes there are mathematical terms that children have never encountered before. Even when we explain things in a simple manner, using language like we are using right

now, some children may still struggle to understand. Therefore, it is necessary to talk about how assessment should be conducted in these geographical locations in undergraduate education.

P4, on the other hand, made the following recommendation about students who may have different needs in a classroom:

When I imagine myself as a teacher, the most challenging aspect for me would be dealing with students with disabilities and gifted students. I haven't quite figured out how to assess them effectively within a classroom setting. If there had been a course specifically addressing inclusive education and gifted students, I believe that current teachers (her peers) would be more knowledgeable in this regard.

A participant (T2) recommended that the teacher education program should teach prospective teachers how to interpret different solutions for the same assessment tasks. She said, *"I believe we need to talk more about how we should interpret two different results of the same evaluation in the lessons."*

Gaining knowledge on how to connect classroom management, lesson planning, and assessment is a recommendation of one of the participants (P4). P4 expressed concerns about making appropriate assessments for students at different levels in a classroom setting while also considering classroom management. She gave an example by mentioning the presence of gifted children in a class who might find the assessments too simplistic and therefore become bored. P4 described her suggestion, as demonstrated below: *"If I could establish a better connection between classroom management and assessment, it would be clearer for me how to assess what, according to which student, and based on their specific needs."*

One of the teachers, T4, emphasized the significance of school administration and its role in the assessment process. T4 clarified his recommendation as follows:

Did we learn about the role of administration in assessment during undergraduate education? I cannot remember, but it should be explained. The importance of the school administration's influence should also be elucidated in relation to this assessment matter.

In conclusion, pre-service and in-service middle school mathematics teachers provided suggestions for enhancing assessment literacy in a teacher education program regarding content. These suggestions included the practical application of assessment knowledge and skills, the structure and sequencing of assessment-related courses in the teacher education program, and enriching the content of assessment courses in line with assessment challenges and daily assessment practices.

4.4.2 How: Method to Improve Assessment Literacy

In addition to the suggestions of the participants to make the education given in the teacher education program more effective in terms of content, they also had some suggestions about how it should be. The analysis of data revealed that the participants' suggestions regarding the "how" aspect can be grouped under three main categories: (i) building a collaborative learning community, (ii) advocating active engagement, and (iii) providing more learning opportunities regarding assessment.

Table 4.11: Suggestions For Enhancing Assessment Literacy in the Method of Teacher Education Program

Main Category	**Sub-category**	**Frequency**
Building a collaborative learning community	Collaboration with experienced teachers including the mentor teacher	10
	Collaboration with the university faculty members	5
	Collaboration with peers	5
	Collaboration with experts	4
Advocating active engagement	Encouraging teacher candidates to actively participate in the assessment processes by taking on more responsibilities during their internship	10
	More student-centered assessment course	6
	Providing pre-service teachers attend parent meetings	4
	Back-to-back teaching experience	3
	Involving pre-service teachers in applying the entire assessment process to real students in an interconnected way like a long-term project	3
Providing more opportunities regarding assessment	Participation in seminars, training sessions, workshops, panels, and conferences focused on assessment	8
	Engaging in independent research and readings on assessment	6

4.4.2.1 Building a Collaborative Learning Community

One of the methods many participants have emphasized for enhancing assessment literacy is the establishment of a collaborative learning community. They stated that collaborative work could improve assessment literacy by creating a learning community. Participants' suggestions in this theme were grouped under four

sub-categories: (i) collaboration with experienced teachers including the mentor teacher, (i) collaboration with the university faculty members, (iii) collaboration with peers, and (iv) collaboration with experts.

All participants dwelled on the significance of collaboration with experienced teachers, including the mentor teacher. They highlighted the importance of exchanging knowledge and expertise between experienced teachers and teacher candidates to enhance assessment literacy through practical wisdom and insights. According to the participants, pre-service teachers can benefit from meeting experienced teachers and learn from their assessment experiences and methodologies. Additionally, they suggested that prospective teachers can enhance their assessment practices by collaborating closely with mentors or experienced teachers during internship courses by observing their ways of managing assessment and providing feedback. For example, T1 said: *"Teacher candidates should be brought together more frequently with experienced teachers. In other words, I think it is needed to keep the experienced and the inexperienced together and benefit from that experience."*

T2, similar to T1, argued that mentor teachers should guide pre-service teachers on assessment. Therefore, T2 supported T1 with the following words:

We have a mentor teacher during the internship course, but we do not exchange information on how that teacher evaluates or gives the final grade. Perhaps, our mentor teachers during the internship can involve us in these aspects. It would be more effective in terms of assessment for pre-service teachers to collaborate with teachers regarding how they grade students, make final decisions, and determine the final grade on the report card.

Another collaboration emphasized by multiple participants was cooperation with university faculty members. Participants emphasized the significance of receiving comprehensive feedback from university faculty members regarding the assessment component of their lesson plans or assignments. They also suggested that university faculty members sharing their evaluation process, including their own experiences, would be beneficial. T5 made the following suggestion regarding the feedback of university faculty members to teacher candidates about their own assessments:

I can suggest receiving feedback from my university professors based on the assessments I have prepared myself. For instance, if I have prepared a very good assessment, I can receive feedback like "Your assessment is very well done." Alternatively, if I have prepared an assessment but there seems to be a certain deficiency, I can seek feedback on that. Receiving constructive feedback of this nature is also a valuable experience.

P1 stated that university faculty members act as role models for teacher candidates, and their assessment methods serve as guiding principles for them. In this regard, P1 has proposed the following suggestion to university faculty members to explain assessment methods and processes:

For example, if our university professors had shown us how they graded and interpreted our midterm results, it would have been much more memorable for me. It would have been a direct and current assessment example for me to understand why I received a BA instead of an AA or the factors my undergraduate professor considered when giving me a BA grade.

Five participants mentioned that collaborating with peers can be a method for improving assessment literacy. They emphasized the benefits of pre-service teachers exchanging ideas and insights with one another regarding assessment. The participants highlighted the importance of the teacher education program facilitating opportunities for group work or discussions. P2 elaborated on her proposal for peer collaboration and outlined how this could be implemented as follows:

We could have consulted each other regarding the findings of our assessments. For example, we could have discussed how to assess, what should be our scoring method, what are the limitations of the methods we use, or whether our assessments are appropriate, and so on. All of these topics could have been discussed among teacher candidates. Individually, let's say I prepared an assessment method and scored accordingly, but it is quite limited. If we had collectively engaged in such idea development, it would have been much more beneficial.

The last suggestion offered by the participants regarding creating a collaborative learning community was a collaboration with experts. Four participants

recommended initiatives such as bringing assessment experts and teacher candidates together in collaboration with the Ministry of National Education (MEB). They mentioned they could enhance their assessment skills by benefiting from these professionals' expertise. For example, T4 expressed his thoughts with the subsequent words:

Experts from the Ministry of National Education or assessment and evaluation specialists can provide seminars on assessment to education faculties. Alternatively, arrangements can be made to bring them together with teacher candidates. So, opportunities can be created for teacher candidates to engage in knowledge exchange with experts in the field of assessment.

4.4.2.2 Advocating Active Engagement

Another method that participants believed would contribute to the development of assessment literacy is advocating for active engagement. Participants mentioned the importance of creating environments or opportunities where teacher candidates can make assessment-related decisions and implement them as experts rather than being treated as novices in the teacher education program. In other words, they emphasized the significance of teacher candidates gaining hands-on experiences where they actively engage in assessment rather than passively receiving information about assessment. The suggestions offered by the participants in this regard were categorized into the following subcategories: (i) encouraging teacher candidates to actively participate in the assessment process by taking on more responsibilities during their internship, (ii) implementing all assessment processes to the real student(s) in a way that is linked like a long -term project, (iii) more student-centered assessment course, (iv) back-to-back teaching experience and (v) allowing pre-service teachers to attend parent meetings.

All of the participants stated that teacher candidates should take more responsibility and participate actively in the assessment processes during their internship. As mentioned in the previous section, the participants highlighted that during their teaching experience in the internship, they were more involved in observation rather than actively participating in the teaching process, and they acknowledged this as

a limitation. They emphasized that the internship is a valuable opportunity for teacher candidates to interact with students and apply their theoretical knowledge of assessment in a teaching experience. They mentioned that this opportunity could be transformed into a chance for them to enhance their knowledge and skills in assessment. T1 pointed out this suggestion as follows:

I wish we could have been more involved in the assessment process during the internship. I would have liked to feel like a real teacher rather than just an intern. It would have been great if we were consulted on matters like determining a student's grade or evaluating their in-class performance. It could have been beneficial if we were involved in the preparation of exam questions, contributing to different areas, but unfortunately, the internship was quite limited. Additionally, it would have been better if we knew that the mentor teacher would involve us in the evaluation process when we started the internship. It would have allowed us to be better prepared and, as a result, more productive.

P5 offered similar suggestions by underlying the importance of giving full responsibility to pre-service teachers in assessing students' mathematical learning:

It might sound like a somewhat utopian example, but during our internships, it could be a process where pre-service teachers are responsible for all the measurement and assessment processes of the students. At the end of the course, they would give the students a final overall grade based on their performance. This process could be developed with the coordination of the internship program and the supervision of education authorities. It could be beneficial to implement such a process in practice rather than just in theory, as it may lead to a more lasting knowledge acquisition, at least from my perspective.

Many participants suggested that the assessment course should be more student-centered. Participants have recommended that the course be designed in a way that encourages teacher candidates to be more actively involved in the class and engage in discussions. They have emphasized that it is challenging to develop a deep understanding of assessment in a course where only university faculty members lecture without active student participation. P3's statements illustrate what these participants recommended:

I think it would be better for our learning if there was more emphasis on practical application and student-centered teaching in this regard, where students are more interactive. Because when only our teachers lecture, the theoretical information doesn't fully come to life in our minds.

Four of the participants recommended that pre-service teachers should attend parent meetings to have opportunities for communication with parents regarding students' assessment results. Therefore, they emphasized that participating in parent meetings, where they can witness the teacher-parent relationship, would provide them with a broader understanding and skills in this aspect. Regarding this, T4 expressed: *"Even though we weren't their own teachers, as teacher candidates, we could sometimes communicate with students' parents. I won't say substitute teacher [like the second teacher of the class], but we could be like the child's second teacher. Or we could attend parent meetings."*

Some of the participants have suggested undergraduate education should provide back-to-back teaching experiences for teacher candidates. The participants mentioned in the deficiencies part that they prepared lesson plans during their undergraduate education but felt lacking in planning the next lesson based on assessment results. In this regard, they recommended that teacher education programs should encourage teacher candidates to prepare and deliver consecutive lesson plans, considering the assessment results. P3 expressed her thoughts on this issue as follows:

I wish we had the opportunity to teach two or more consecutive lessons during our internship instead of just one. For example, let's say I'm teaching multiplication with fractions. This topic cannot be covered in just one lesson. If I had been given the chance (as a teacher candidate) to teach the entire topic and at least identify the students' conceptual misconceptions and gaps, I could have prepared activities related to those issues for the second lesson. This way, I would have created a plan to address these conceptual misconceptions. In short, I think it would have been beneficial to have more opportunities for teaching multiple lessons.

Another suggestion from the participants was to involve pre-service teachers in applying the entire assessment process to real students in an interconnected way like a long-term project. Participants mentioned that by designing a project in which

teacher candidates take charge of the entire process, from start to finish, they can gain valuable long-term experience. For instance, P5's recommendation was:

It could be a real-life project for teacher candidates. They would be responsible for designing and managing the entire assessment process, and after that, they would extract the necessary assessment data. They could even write a small-scale paper related to this. It could be envisioned as a long-term assignment, like a term project.

4.4.2.3 Providing More Learning Opportunities Regarding Assessment

Another method that participants believed would enhance teacher candidates' assessment literacy is for the teacher education program to provide more opportunities related to assessment. Accordingly, the participants' recommendations were categorized as (i) participation in seminars, training sessions, workshops, panels, and conferences focused on assessment, and (ii) engaging in independent research and readings on assessment.

Many of the participants suggested that the teacher education program should provide opportunities for teacher candidates to enhance themselves through assessment-focused seminars, training sessions, workshops, panels, and conferences. They mentioned that this would allow them to continue their professional development beyond the courses in their undergraduate education. In line with this, T2 expressed the following:

Conferences can be organized. Seminars can be held. For instance, it could be focused on preparing a test specifically for assessment and evaluation. If there won't be a mandatory course or elective course, opportunities can still be provided in the form of workshops. Activities that support teacher candidates can be conducted in this regard.

Another suggestion of the participants is to encourage teacher candidates of the teacher education program to conduct independent research and readings on assessment. T3 expressed his request in these words:

Well, we really need to understand what's happening in the world individually.

We get trapped in a very small environment. In this sense, it is essential to look internationally. That's why it is necessary to do reading.

To sum up, pre-service and in-service middle school mathematics teachers suggested that building a collaborative learning community, advocating active engagement and providing more learning opportunities regarding assessment would be effective strategies to enhance assessment literacy.

4.4.3 Why: Orientation to Improve Assessment Literacy

One of the suggestions provided by the participants concerned the orientation of teacher education programs. According to Juanjuan and Mohd Yusoff (2022), the "why (orientation)" theme emphasizes that improvement programs should focus on helping teachers reconsider their role as assessors and supporting them in becoming reflective practitioners. Participants made only one suggestion in this regard, which was categorized as cultivating reflective practitioners, similar to the study by Juanjuan and Mohd Yusoff (2022).

Table 4.12: Suggestions For Enhancing Assessment Literacy in the Orientation of Teacher Education Program

Main Category	Sub-category	Frequency
Cultivating reflective practitioner	Teacher education programs should emphasize the importance of assessment to a greater extent	2

4.4.3.1 Cultivating Reflective Practitioner

Two participants emphasized the significant role of assessment in the teaching and learning process and expressed that this importance is not adequately reflected in teacher candidates. They highlighted the need for teacher education programs to make teacher candidates more aware of the importance of assessment. For example, T4 said:

Teacher candidates should be better acquainted with the importance of assessment.

In short, participants recommended that teacher education programs should emphasize the significance of assessment for learning more and train teachers who value assessment.

CHAPTER 5

DISCUSSION, IMPLICATIONS AND RECOMMENDATIONS

One of the purposes of the study was to investigate the assessment literacy levels of pre-service and in-service middle school mathematics teachers. The study also aimed to obtain the participants' thoughts about the deficiencies in the undergraduate education they received regarding assessment literacy and their suggestions for improving the quality of assessment education in undergraduate education. In this section, the study's main findings will be presented and discussed in detail. In addition, conclusions will be provided by discussing previous studies. Finally, the implications and recommendations for further research will be examined considering the present study's findings and discussion.

5.1 Assessment Literacy Levels of Pre-service and In-service Teachers

According to the findings of the study, the average number of correct answers out of 30 questions for pre-service teachers is 15.93 (53.1%), while for in-service teachers it is 15.73 (52.4%). Yamtim and Wongwanich (2014) classified assessment literacy scores into three levels: inadequate or in need of improvement (below 60%), satisfactory (60%-79%), and excellent (80% and above). Accordingly, descriptive analysis of the quantitative data of the current study revealed that the assessment literacy level of pre-service and in-service middle school mathematics teachers were low. This finding aligns with previous studies that have reported a lack of assessment literacy among both pre-service and in-service teachers (Akayuure, 2021; Alkharusi et al., 2011; Beziat & Coleman, 2015; Black, 2002; Brookhart, 2001; Dehqan & Asadian Sorkhi, 2020; DeLuca & Klinger,

2010; Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk, 2014; Plake et al., 1993; Popham, 2004; Siegel & Wissehr, 2011; Stiggins , 1995; Volante & Fazio, 2007; Yamtim & Wongwanich, 2014). According to DeLuca and Bellara (2013) insufficient undergraduate assessment education, which is heavily theoretical, detached from teachers' day-to-day assessment practices, and possibly mismatched with contemporary educational assessment standards, might be a reason for the low levels of assessment literacy. These reasons might also be applicable to this study because the analysis of data indicated that the participants reported similar factors as the shortcomings of the teacher education program.

The independent t-test conducted to determine if there was a significant difference in the assessment literacy levels between the middle school mathematics pre-service and the in-service teachers who participated in the study revealed no significant difference in the participants' scores. The literature is inconsistent regarding whether pre-service or in-service teachers have higher assessment literacy. Some previous research findings have shown that in-service teachers have higher levels of assessment literacy compared to pre-service teachers (Dehqan & Asadian Sorkhi, 2020; Mertler & Campbell, 2005). Researchers argued that in-service teachers are more assessment literate than pre-service teachers because of their experience of applying the learned assessment theoretical knowledge in real classroom settings. For some other studies, pre-service teachers have higher assessment literacy levels than in-service teachers (Alkharusi et al., 2011; Çambay, 2016. According to Alkharusi et al. (2011), the reason for the higher level of assessment literacy among teacher candidates compared to in-service teachers may be their more recent fulfillment of assessment courses. In other words, pre-service teachers may have achieved a higher level of assessment literacy than in-service teachers because they remember more theoretical information about assessment. However, these differing results in the studies may suggest that the assessment education has deficiencies in both theoretical and practical aspects. This idea is also supported by the quantitative finding of the current study; there is no significant difference between the assessment literacy levels of pre-service and in-service teachers. The absence of a significant difference in assessment literacy scores could imply that both pre-service and in-service teachers might have similar gaps or limitations in their understanding of assessment concepts, techniques,

and practices. In other words, they might share common areas of weakness or misconceptions related to assessment. This finding also suggests that teacher education programs, both for pre-service and in-service teachers, may not adequately address the development of assessment literacy. This implies that these programs might need to reevaluate and enhance their curriculum to better equip teachers with the necessary assessment knowledge and skills. It may call into question whether the subject matter covered in teacher training courses aligns with classroom instructors' real-world requirements. Additionally, if pre-service and in-service teachers have similar assessment literacy levels, it could indicate a need for more alignment between what is being taught in these programs and what teachers encounter in their classrooms.

In the study's quantitative analysis, each competency area that constituted the assessment literacy inventory was examined separately for pre-service and in-service teachers. It was determined that the lowest average performance demonstrated by middle school mathematics pre-service teachers, according to the competency areas of assessment literacy, was in the fifth competency area, "creating accurate techniques for grading students based on their assessments" (M=1.85), while the highest average performance was in the third competency area, "administering, scoring, and interpreting the assessment results" (M=3.35). The lowest average performance demonstrated by middle school mathematics in-service teachers, according to the competency areas of assessment literacy, is the same as pre-service teachers, and it is in the fifth competency area, "creating accurate techniques for grading students based on their assessments" (M=1.78). The highest average performance is in the first competency area, "selecting proper assessment techniques, tools, and methods for teaching decisions" (M=3.32). In other studies, it has been observed that both pre-service and in-service teachers obtain low scores in the fifth competence area (Akayuure, 2021; Beziat & Coleman, 2015; Dehqan & Asadian Sorkhi, 2020; Mertler & Campbell, 2005; Yamtim & Wongwanich, 2014). It can be said that there are significant areas that need improvement in teacher education programs in the fifth competence area, which involves creating accurate techniques for grading students based on their assessments considering the findings of other studies. Pre-service teachers obtaining the highest score in the "administering, scoring, and interpreting

the assessment results" competency area might indicate that they have been exposed to more detailed theoretical knowledge about this subject during their undergraduate education. Alternatively, it could be due to having had more opportunities for practical application in this area. In addition, the fact that in-service teachers attained the highest score in the "selecting proper assessment techniques, tools, and methods for teaching decisions" competency area could suggest that they have learned a broader range of assessment methods during their undergraduate education and are actively implementing them.

5.2 Deficiencies of the Teacher Education Program in Developing Assessment Literacy

In this study, interviews were conducted with ten participants after administering the assessment literacy inventory, considering their assessment literacy levels and other factors, as explained in the methodology section. In these interviews, pre-service and in-service middle school mathematics teachers were asked specific questions about the teacher education program's deficiencies in assessment literacy for each competence area. The deficiencies mentioned by the participants were categorized based on the theoretical and practical deficiencies of the teacher education program. Therefore, the shortcomings of undergraduate education for each competence area will be discussed from both theoretical and practical perspectives in this chapter. Furthermore, this section will discuss the disparity between theoretical knowledge learned in teacher education programs and real classroom applications, as evident in both the literature and the current study's findings. As a result, insufficient theoretical foundations offered in the teacher education program will be presented first, followed by the practical shortcomings and, ultimately, the issue of the gap between theory and practice in the teacher education program. These will be presented in consideration of each competency area in this section. These deficiencies will be explored through the conclusions described in the preceding section.

5.2.1 Deficiencies of Theoretical Knowledge

Insufficient assessment knowledge and skill is one of the reasons why teachers and teacher candidates have low levels of assessment literacy (Coombe et al., 2012; Dehqan & Asadian Sorkhi, 2020; Mertler & Campbell, 2005; Popham, 2004). The findings of the current study indicated that participants reported theoretical deficiencies in the assessment courses. In order to train teachers who are literate in assessment, teacher education programs should offer them essential knowledge and skills in assessment (DeLuca & Klinger, 2010). The assessment education provided by teacher education programs should be based on a rich foundation that encompasses the necessary theoretical knowledge and aligns with assessment standards (DeLuca et al., 2016).

The analysis of data revealed the areas that need improvement in undergraduate education specific to each competency area. Some of the deficiencies mentioned by the participants were categorized under the title of general theoretical deficiencies. One of the general theoretical deficiencies mentioned by a significant number of participants and applied to multiple competency areas was the superficial content of the assessment-related courses. The participants expressed that they felt ill-equipped in terms of theoretical knowledge to deal with different situations that they may encounter in real classroom environments and discussed the powerlessness they felt in knowing what to do in such situations. For instance, they indicated that in the teacher education program, they learned that an assessment should be clear and fair for all students in a class. However, they pointed out that specific methods and theoretical foundations of them on how to achieve this in assessment courses were not discussed. Accordingly, it can be concluded that the according to the participants assessment courses are insufficient in providing comprehensive theoretical information about the diverse needs that teacher candidates may encounter while conducting evaluations in a real classroom environment. This finding aligns with the study by DeLuca and Klinger (2010), who also highlighted teacher candidates' inadequate acquisition of knowledge in assessment courses.

Another deficiency highlighted by both pre-service and in-service middle school mathematics teachers is the limited availability of elective and mandatory courses

on assessment. As DeLuca and McEwen (2007) also noted, despite the importance of assessment as an integral part of education and a skill teachers should possess, assessment courses are not even mandatory in some universities' teacher education programs. The small number of assessment-related elective and mandatory courses may account for also the lack of depth in assessment-related courses. As there was only one compulsory assessment course in the teacher education program in which participants were enrolled, it would be difficult to go into the details of many assessment principles.

Another shortcoming mentioned by the participants was the inability to thoroughly grasp the assessment concepts because they took the assessment course too early. The first four semesters of the teacher education program offered more mathematics courses than education courses. The participants in the study took the assessment course in the fourth or fifth semester. Therefore, they stated that when they took the assessment course, their knowledge and skills related to mathematics education were underdeveloped, and they were not very conscious. Therefore, they remarked that this led to their incomplete understanding of assessment concepts and principles.

The theoretical deficiencies of the teacher education program stated by the participants for the 1st and 2nd competency areas were lack of knowledge in choosing/developing an assessment that can meet the diverse classroom needs, inadequate knowledge on assessment preparation procedure and process, lack of knowledge in developing an assessment that attends to the educational and contextual difficulties, and lack of readiness to handle language-related challenges in assessments. The biggest deficiency mentioned by the participants was the challenge of selecting or designing assessments that would cater to different needs, such as students with varying academic levels (high-medium-low) or students with different abilities, including gifted and disabled students in a class. Pre-service teachers expressed their lack of knowledge about choosing suitable assessments for students with diverse needs, while in-service teachers also voiced the difficulty in selecting appropriate assessments due to inadequate undergraduate education on this topic. Moreover, they stated that teacher training needed to be improved, especially in guiding them on how to design a summative assessment process. For example, they mentioned not knowing how to determine the number of questions from each topic in

a written exam and how much weight each question should carry in the overall grade. Along the same lines, the literature (e.g., Norman (2010)) remarks that designing successful, reliable, and valid assessments for students with diverse backgrounds, learning styles, and cultures is challenging for teachers.

Lastly, a teacher in a rural school mentioned facing language barriers and feeling that the assessment knowledge gained during undergraduate education was insufficient. Considering these study findings, a more profound and broader perspective education should be offered for real classroom needs in selecting and developing appropriate assessment methods in teacher education programs. In line with this finding, Ladson-Billings (2009) also highlighted that students in rural areas have distinct needs, which makes it challenging for teachers to consider these diverse contexts when conducting assessments.

The theoretical deficiencies in the 3rd competency area were related to scoring and interpreting the assessment results. Participants noted that the assessment-related courses, including statistics courses, were insufficient to help interpret assessment results and identify students' difficulties or misconceptions. Similarly, Lazer (2009) also stated that teachers find the statistics that enable them to make comments about student performances and reveal the relationship between student understanding and teaching effectiveness to be complicated and challenging to use. Another aspect of the teacher education program in this area of competence is the scoring of the assessment results. The participants underlined that the sample rubrics used in the teacher education courses offered by their faculties were superficial and therefore did not add to deep knowledge. For example, they remarked that they did not know what to do when an unexpected answer or solution came from the students. This finding overlapped with some other studies, which reported teachers' difficulties in preparing scoring rubrics and interpreting assessment results (Andrade, 1997; Uludag & McDonough, 2022).

There were two main theoretical shortcomings reported by the participants regarding the teacher education program in the 4th competence area (making decisions about students, planning the teaching, and developing the curriculum): lack of knowledge about taking individual differences into account when making decisions based on

assessment results and lack of addressing what can be done to improve teaching after an assessment is applied. The reason for these deficiencies in this regard could stem from both observation and application shortcomings in lesson planning based on assessment outcomes during the teacher candidates' undergraduate education. The analysis of results showed that although the participants were aware of the need to make teaching decisions based on assessment results, they lacked the knowledge of how to effectively manage this process. A similar findings were found in other studies in the literature (Boikhutso, 2010; Sugianto, 2020; Yuan & Zhang, 2016) which revealed that teachers were unsuccessful in effectively utilizing assessment results while preparing lesson plans. These findings supported that teachers might not have learned how to effectively integrate assessment into lesson planning within their teacher education program.

In the 5th competency area, namely developing valid pupil grading procedures, the participants stated that the teacher education program had significant theoretical deficiencies in equipping them with the necessary information for developing valid grading systems/procedures. The teacher participants also noted that the teacher education program failed to present the facts of the Turkish education and evaluation system for valid grading, for example, passing a student who should not have progressed to the next grade level considering the assessment results. In the literature, many studies (Brookhart, 2001; Cizek, 2010; T. R. Guskey & Bailey, 2001; McMillan, 2003; McMillan & Nash, 2000; Sadler, 1998), concluded that teachers and teacher candidates were not prepared to adequately perform valid grading procedures.

Finally, the participants commented that the teacher education program had theoretical deficiencies in helping them communicate the assessment results to students, parents, and other educators (the 6th competency area), particularly, the knowledge to effectively report children's low grades to their families and help them accept it. The participants stated that they did not learn how to use official reporting methods when communicating about assessment results, and they emphasized that this should be taught to pre-service teachers during their undergraduate education. This finding of the present study is aligned with the study by Eret (2013). According to the Eret (2013), teacher education fails to sufficiently equip prospective teachers to effectively communicate with parents and colleagues within the school setting,

address the needs of students with special education requirements, and adapt to various teaching environments. Similarly, Brookhart (1999) indicated that teacher education fell short in terms of facilitating effective communication about assessment outcomes, emphasizing the need for more comprehensive courses in this regard.

5.2.2 Deficiencies in Practice/Experience

In addition to theoretical knowledge, putting theoretical knowledge into practice also plays an essential role in educating teacher candidates in a more pedagogical way (Janssen et al., 2015). As practice encompasses knowledge, identity, evaluation, and abilities, gaining proficiency in such a fundamental practice involves acquiring a deeper theoretical understanding of its significance and when it can effectively promote students' learning (Grossman & McDonald, 2008). Similarly, Avalos (2011) remarked that it is critical for teachers to have teaching practice as a central aspect of their professional development and to build an understanding of applying their knowledge in a way that encourages students' improvement. Otero (2006) claimed that implementing assessment methods is more challenging for teachers with no experience. Although practice is an important part of teacher education, many studies have revealed that both teacher candidates and teachers were unable to apply their assessment knowledge in actual assessment practices (Ogan-Bekiroglu & Suzuk, 2014). One of the main reasons for this is that teacher candidates do not have enough opportunities to apply their theoretical knowledge in teacher education programs (Dehqan & Asadian Sorkhi, 2020; Essomba, 2010; Mertler & Campbell, 2005; Zembal-Saul et al., 1999).

The current study's findings a also are also in line with the literature by demonstrating the practical areas that need improvement in teacher education programs. Similar to the theoretical shortcomings, there were specific practical deficiencies in each competence area, but some were also mentioned for multiple competence areas. These practical deficiencies were also analyzed and categorized under general practice/experience deficiencies.

One of the overarching practical deficiencies is putting the assessment into the background and not giving enough emphasis on it in the teaching experience.

Teaching experience allows teacher candidates to apply the theoretical knowledge they have learned in their undergraduate education and develop themselves. However, most participants mentioned that during their teaching experience (internship), they prioritized aspects such as how to deliver the lesson, what materials to use, and how to manage the classroom, which led them to put the assessment into the background. Kaplan (2015) conducted a study and found that teacher candidates used formative assessment methods very little or not at all to assess students in their lesson planning. Additionally, the researcher asserted that most participants overlooked the fact that feedback from the assessment section of lesson plans could also benefit students. In parallel, Yilmaz-Tuzun (2008) stated in her study that teacher candidates had theoretical knowledge about assessment but did not integrate it into practice. Accordingly, these findings of the study overlap with other studies. The reason for teacher candidates putting the assessment into the background in practice may be that they have not fully grasped the importance of assessment or that their mentor teachers or faculty instructors also neglect assessment.

Another general practical deficiency expressed by the participants was a lack of getting feedback on assessment parts of tasks. For instance, the participants mentioned that they obtained very little feedback on the assessment part when they received feedback from faculty instructors on the lesson plans that they prepared. Quality feedback is important for student's growth and improvement (Carless et al., 2011). Therefore, offering high-quality feedback is essential for successful instruction and a critical criterion for achieving students' anticipations (Ramsden, 2003).

Although feedback is critical for effective learning, the participants stated that they received very little feedback from university faculty members related to assessment. The reason for this might be the same as the previous deficiency, which is the insufficient importance and attention given to assessment by university faculty members. The low diversity of schools that they attended in teaching experience was another shortcoming said by the participants. They stated that the schools where they did their internship had a similar culture and student profile, giving them a limited perspective. For example, they claimed that their teaching experience was mainly in schools that generally performed above average. As a result, they could

not observe students in schools with lower performance levels or in more challenging environments. Another of the shortcomings was that the internship experience fell short in terms of observing the assessment process used by the mentor teacher. The participants mentioned that they could not observe the assessment methods used by the teachers who mentored them during the internship course. According to the participants, one of the reasons for this is that the mentor teacher does not use enough assessment, and another reason is that the teacher candidates are not given enough opportunity to observe formal assessment processes such as scoring assessments or planning lessons towards assessment. In parallel with the findings of the current study, Mohono-Mahlatsi and Van Tonder (2006) concluded in their research that teacher candidates lacked sufficient feedback and opportunities for implementation regarding assessment from mentor teachers during their teaching experience. According to these findings, it can be said that the internship experience is lacking in observing the assessment methods of mentor teachers who are more experienced than teacher candidates and making inferences accordingly. The assessment course was not student-centered was another shortcoming expressed by the participants. One of the important factors ensuring students' optimal learning in classes is student-centered courses, where students actively participate (Emaliana, 2017). However, despite this, most courses are teacher-centered (Estes, 2004). Participants asserted that the lessons typically moved through slides and primarily focused on the teacher. They emphasized the absence of teacher-student interaction in the assessment course and said that the teacher candidates' level of participation in the subject was poor. They emphasized that due to the inability to participate in the assessment course actively, their interest in the class was low, and they felt that the class was ineffective.

Lastly, the participants declared the issue of late opportunity to practice in teaching experience after learning the theoretical knowledge of undergraduate education. As mentioned in the theoretical deficiency, the participants stated that they took the assessment course early, making it challenging to grasp the acquired knowledge. Furthermore, they also mentioned that they took the internship course in the last two semesters, and there was a significant time gap between the assessment and internship courses. They emphasized that this situation made it challenging to apply the assessment knowledge during their internship. Considering the finding, it can

be said that the considerable time gap between the assessment and internship course where teaching experience is gained makes it difficult for teacher candidates to apply the knowledge they have learned.

The practical deficiencies of the teacher education program stated by the participants for the 1st and 2nd competency areas were the limited exposure to a variety of assessment methods and tools/tasks on assessment. Participants mentioned that pre-service teachers consistently use the same assessment methods in tasks, such as lesson plans, prepared as part of the teacher education program. Although they attempted to incorporate different activities or teaching methods in these tasks, they stated that they always remained stuck in the same cycle in terms of using different assessment methods. Consequently, they highlighted the lack of opportunities to observe diverse assessment methods in undergraduate education as a shortcoming. Regarding this finding, it can be said that pre-service teachers were not appropriately encouraged to use different assessment methods. In line with this, Rey et al. (2012) also declared that teacher candidates might encounter difficulties in effectively integrating various fundamental procedures within the appropriate context, choosing the suitable ones within a structured assignment, and grasping primary procedures individually.

The practical deficiencies of the 3rd competency area were the insufficient practice of newly acquired statistical concepts and the lack of experience in applying a rubric to assess actual student work. Participants mentioned that after learning terms such as standard deviation and z-tables in the statistics course, they did not have the opportunity to solve questions and apply this knowledge. For this reason, they stated that the undergraduate education was insufficient to make sense of these statistical concepts that they should use while evaluating. These findings are confirmed by the study conducted by Parsian and Rejali (2011), which aimed to gain insights into the preparation of mathematics teachers for teaching statistics in Iran. The researchers concluded that teacher education programs fell short in providing adequate statistical training both in terms of theoretical understanding and practical application. They also noted that although they had experience preparing rubrics, they did not apply them to real students. For this reason, they expressed that they could not test the quality and suitability of the rubrics they prepared. Similarly, Luft (1999) aimed

to observe the contribution of a secondary science methods course to prospective science teachers' understanding of rubrics. In his study, he concluded that prospective teachers should be exposed to more rubric applications during their education and that teacher preparation programs should offer more opportunities for practical implementation. These findings may indicate that the teacher education program is limited in terms of providing teacher candidates with opportunities to apply their theoretical knowledge on real students.

Most participants noted insufficient practice in planning the subsequent lesson according to the assessment results in the practical deficiency of the 4th competency area. According to Richards and Renandya (2002), one of the most essential and critical points teachers should consider in lesson planning is students' assessment results. With assessment results being crucial in lesson development, the lack of opportunity for teacher candidates in teacher education programs to prepare a lesson plan by using assessment outcomes can be considered a significant deficiency. The inability of pre-service teachers to gain this experience may cause them to encounter difficulties in planning lessons using evaluation results once they become teachers. The findings of this study are in line with the existing literature. For instance, Campbell and Evans (2000) examined the assessment practices of teacher candidates who had successfully completed an assessment course by analyzing their lesson plans. According to the results of their research, numerous teacher candidates did not effectively integrate assessment into their lesson plans. The researchers concluded that the limited opportunities for teacher candidates to practice integrating assessment into their lessons during the assessment course contributed to the inconsistency between assessment instruction and practical application among teacher candidates. They indicated the need for further research to explore the factors contributing to this gap.

Lastly, in the 6th competency area, the deficiency that the participants stated about the lack of practice of the teacher education program was communication difficulties due to the perceptions of the students in the internship school toward pre-service teachers. Participants have shared that, even though they had the opportunity to gain experience in communicating with students during the internship course, the students did not perceive the teacher candidates as their teachers. They expressed that students saw

them as sister-brother rather than a teacher, so they did not take them very seriously when communicating with them about the assessment results. When considering this finding, it can be said that mentor teachers should involve teacher candidates more like actual teachers in the classroom during internship courses.

5.2.3 Gaps Between Theory and Practice

The majority of the participants stated that the theoretical education provided by the teacher education program regarding assessment was not in line with practical application. Especially all of the in-service teachers mentioned that there was a significant gap between theory and practice in assessment, leading to difficulties in conducting assessments. For example, they expressed that the knowledge gained during their undergraduate education concerning classroom environments, student levels, school culture, and student needs was founded on ideal scenarios, yet the actual real-life situations differed significantly from what they had learned in theory. Concerning this, Rasmussen and Rash-Christensen (2015) concluded that teacher education programs fall short of adequately preparing students for actual classroom settings. In their study with physics prospective teachers, Ogan-Bekiroglu and Suzuk (2014) also found that they faced considerable challenges when trying to put the theoretical foundational knowledge they had gained in assessment literacy into practical application. In their research, Siegel and Wissehr (2011) discovered considerable discrepancies between teacher candidates' theoretical conceptions of assessment as expressed in their teaching philosophies and journals and their actual experiences of assessing students in the lesson plans that they created for instruction. Furthermore, numerous different studies have also revealed a gap between theory and practice in teacher education programs (DeLuca & Klinger, 2010; DePaul, 2000; Husebo, 2012; Korthagen, 2010; Krumsvik & Smith, 2009; Schwab et al., 2021; Yilmaz-Tuzun, 2008). Therefore, the findings of this study are consistent with the existing literature. This gap between theory and practice in assessment may arise from teacher education programs not adequately preparing teacher candidates for real-life situations. In other words, as participants also pointed out, the theoretical knowledge provided in assessment education may be primarily focused on ideal classroom conditions.

The participants also stated that insufficient opportunities to apply theoretical knowledge of assessment in practice was another deficiency of the teacher education program. Participants declared that they could not put the theoretical knowledge they learned for almost all competency areas into practice, and they considered this a significant deficiency in teacher training. They cited that they could not find chances to apply their assessment knowledge to real students despite having compulsory internship courses. Alternatively, they indicated that they didn't produce any assessment-related products (such as assessment tools, rubrics, or analyses) using the theoretical knowledge they had gained. They further pointed out this situation left the theoretical knowledge they acquired hanging in the air and hindered their complete understanding of the subject. To train teachers who possess assessment literacy, it is necessary for them to be both theoretically equipped and to apply this knowledge in practice (Stiggins , 1995). Although experiencing and making connections between concepts is crucial for effective learning, teacher education programs fail to provide sufficient opportunities for teacher candidates to practice, resulting in lower levels of assessment literacy among teachers (DeLuca et al., 2013). In their study, Volante and Fazio (2007) administered an assessment literacy questionnaire to prospective teachers for four years. Consistent with the current study's findings, the researchers stated that, based on the survey results, most teacher candidates expressed that the teacher education program provided insufficient opportunities for practicing assessment and that they needed more practical experience in this area. In other words, the literature and the findings of this study are consistent with the lack of assessment practice in teacher education (Arter, 2001; Brookhart, 1999; DeLuca & Klinger, 2010; Ediger, 2000; Mertler, 2009; Ogan-Bekiroglu & Suzuk, 2014). Lack of practice may be one of the reasons why teachers and prospective teachers have low assessment literacy, as the literature agrees. Moreover, these findings may also be one of the reasons for the gap between theory and practice. The fact that pre-service teachers do not have the chance to practice in experiences such as lesson plans or internships during undergraduate education may prevent them from making connections between assessment concepts, and, as a result, their development in assessment literacy may be insufficient.

5.3 Suggestions of the Teacher Education Program in Developing Assessment Literacy

The pre-service and in-service middle school mathematics teachers' suggestions on how undergraduate education can provide a more effective education in assessment literacy. The suggestions provided by the participants were classified into three main themes, similar to the study conducted by Juanjuan and Mohd Yusoff (2022): a) "What" refers to the content of the assessment literacy improvement program for teachers, b) "How" pertains to the approach or method of the assessment literacy improvement program for teachers, and c) "Why" focuses on the underlying principles or orientation of the assessment literacy improvement program for teachers. These suggestions will be investigated further in the following section.

5.3.1 What: Content of a Program to Improve Assessment Literacy

Some of the suggestions made by the pre-service and in-service middle school mathematics teachers to enhance the effectiveness of undergraduate education in terms of assessment literacy were related to the content of the teacher education program. They have emphasized the need for contextualization of assessment-related courses offered in the teacher education program and the design of the program to be closely connected to real assessment practices. These suggestions are practical application of assessment knowledge and skills, the structure and sequencing of assessment related courses in the teacher education program and enriching the content of assessment courses in line with assessment challenges and daily assessment practices. Detailed information on each suggestion is given in the following sections.

5.3.1.1 Practical Application of Assessment Knowledge and Skills

The participants have stated that they will acquire and develop assessment literacy through actual assessment practices. They have expressed the importance of the teacher education program offering chances to apply theoretical knowledge to authentic assessment practices in various contexts. Towards this, almost all

participants pointed out that the courses should incorporate a diverse range of assessment-related assignments for pre-service teachers. The participants have stated that they will further develop their skills in selecting and developing appropriate assessment methods, creating rubrics, analyzing and interpreting assessment results, and scoring or grading by engaging in assignments where they can apply the theoretical knowledge they have acquired in their undergraduate education. Another suggestion offered by the participants was that the courses need to include actual or hypothetical student assessment results. They mentioned that this approach would enhance their understanding of assessment concepts in real-life contexts. For instance, they cited that analyzing sample assessment results and creating a lesson plan based on them would make the evaluation practice more tangible and beneficial. Additionally, the participants mentioned that they had prepared numerous lesson plans during their undergraduate education, but they often neglected assessment in these plans. However, they believe that assessment should be an essential and integral part of lesson planning. Therefore, they suggested that the courses need to place greater emphasis on incorporating the assessment part into lesson planning. They remarked that the active inclusion of assessment in pre-service teachers' lesson plans would not only allow them to gain practicality in selecting and developing appropriate assessments but also help them develop a deeper understanding of how assessment influences instructional decisions and supports students' learning. Lastly, one of the suggestions was that the program should require preservice teachers to practice a wide variety of assessment methods. It was mentioned that in undergraduate education, the same assessment methods are often used in assignments, limiting the opportunity to observe or experience different assessment methods. Therefore, they noted that using various assessment methods in tasks within the teacher education program would broaden their assessment perspectives and knowledge. This finding aligns with the findings of Ayalon and Wilkie (2020). They also have recommended that practicing with sample student solutions enhances the assessment literacy of pre-service teachers, and, therefore, they suggest incorporating activities involving real-school contexts like student solutions in teacher education programs. Gotch and McLean (2019) have made similar suggestions by emphasizing the importance of developing conceptual and theoretical insights regarding the relationship between assessment literacy, contextual variables, and pupil results. Moreover, the recommendations

of Juanjuan and Mohd Yusoff (2022) are consistent with these findings. They also expressed that teachers develop their assessment literacy by participating in assessment activities. Therefore, teachers should be allowed to transform their theoretical knowledge into real-world assessment procedures in certain contexts.

5.3.1.2 The Structure and Sequencing of Assessment Related Courses in the Teacher Education Program

Some of the suggestions made by the participants regarding the content of the teacher education program were related to the structure and sequencing of assessment-related courses in the teacher education program. Many participants mentioned that the number of compulsory and elective courses offered on assessment in the teacher education program was inadequate. As a result, they could not take the courses that would help them improve their assessment skills. Additionally, they mentioned that there are many important subtopics within assessment and that taking a single course on assessment in undergraduate programs is not sufficient to grasp all these topics. For this reason, they have suggested increasing the number of compulsory and elective courses related to assessment. This finding is consistent with several other studies that suggest an increase in the number of assessment-related courses in teacher education programs (DeLuca & McEwen, 2007; Stiggins , 2004; Volante & Fazio, 2007). Another suggestion that pre-service and in-service middle school mathematics teachers mentioned was to consciously take the assessment course in the later stages of the undergraduate program. The participants stated that they took the assessment course at the start of their undergraduate studies and were unable to create a strong foundation of knowledge and abilities in their various fields of study within that period. When the Elementary Mathematics Education (EME) program at the university where the study was conducted was revised in 2018, the compulsory assessment course was moved from the 4th semester to the 5th semester. Therefore, even though the assessment course was taken a semester later according to the updated program, one of the pre-service teachers who took the assessment course based on the new program also made the same recommendation. This may indicate that it would be better to take the assessment course in the later years, after the 6th semester. The last suggestion of the participants regarding the structure and

sequencing of the assessment-related courses in the teacher education program was configuring the program so that the connections between assessment-related concepts learned in different courses can be established. The participants mentioned that they encountered concepts related to assessment in other courses besides the assessment course. However, they highlighted the deficiency of not establishing connections to assessment in those courses. One participant likened the courses in the teacher education program to pieces of the puzzle and emphasized the need to put the pieces of the puzzle together to achieve meaningful learning in assessment. For example, they mentioned that in one of the courses during their undergraduate education, they learned how to teach each topic to students and identify possible misconceptions. In line with this, they suggested that an assessment component could be added to that course at the end of each topic to assess students' understanding. Therefore, it can be said that establishing connections between assessment concepts with the content of other courses offered in teacher education programs can contribute to the assessment literacy of teacher candidates. For this purpose, it is important for faculty members delivering these courses to review the program with a focus on establishing these connections.

5.3.1.3 Enriching The Content of Assessment Courses in Line with Assessment Challenges and Daily Assessment Practices

Enriching the content of assessment courses in line with assessment challenges and daily assessment practices was determined as one of the content suggestions categories. Participants reported several deficiencies related to theoretical knowledge in some aspects of assessment in the teacher education program. Based on these deficiencies, the participants proposed some topics that need to be added to the assessment education in the teacher education program. The most mentioned topic among these was receiving more information about how the grading process should be. Many participants stated that their undergraduate education provided limited knowledge about grading but highlighted the need for more emphasis on this topic. Because they emphasized that, as a final outcome, they need to grade students based on all the assessments conducted, but the teacher education program should provide a richer content on how to do it. Additionally, according to the results of quantitative

data analysis, both pre-service and in-service teachers obtained the lowest score in the competence area of the 5th domain, which is "Creating accurate techniques for grading students based on their assessments." Both quantitative and qualitative data analysis indicated the need for teacher education programs to provide more comprehensive and in-depth knowledge in this area.

The second most frequently mentioned topic by the participants was need for knowledge about how to plan instruction according to different assessment results. They noted that undergraduate education was lacking in terms of planning instruction based on different assessment results from the same assessment or considering multiple assessments, as well as in scenarios where desired assessment outcomes were not achieved. Parallel to these deficiencies, it can be said that there is a need for teacher education programs to provide more information in undergraduate education about how to develop lesson plans in response to different assessment outcomes and the resulting scenarios.

The third suggestion offered by the participants was to customize the theoretical knowledge provided in the courses according to the needs of the students. During their teacher education program, the participants mentioned that they learned how to assess in an ideal classroom environment. However, they also highlighted that real classroom situations may present different needs. For example, they struggled with assessment preparation due to language barriers in rural schools or did not know how to assess students with disabilities or gifted students in a classroom. Therefore, it can be said that teacher education programs need to provide more guidance on how teachers can assess students' specific needs in order to meet the real classroom demands in their professional careers. Participants' fourth suggestion was to gain knowledge about how to interpret different solutions for the same assessment. The in-service participant who offered this suggestion expressed difficulty in interpreting different solutions to the same question in the same assessment. The participant argued that undergraduate education should train teacher candidates to be more conscious and knowledgeable in this regard. In other words, the participants emphasized the need for more detailed and in-depth theoretical knowledge about rubrics.

Another suggestion was about the need for training on how to connect class- room management, lesson planning and assessment. The participants expressed that they lacked sufficient knowledge on how to effectively manage classroom while conducting assessments. In addition, they stated that knowing how to create a lesson plan based on successful classroom management, following the assessments, would be beneficial in selecting the most appropriate assessment method. For this reason, they mentioned that integrating classroom management and lesson planning into assessment during undergraduate education would contribute to the training of more competent teacher candidates.

The last suggestion to enrich the content of assessment courses in line with assessment challenges and daily assessment practices was to gain information about school administration and its role in the assessment process. The participants emphasized the crucial role of administration in the assessment process and highlighted the importance of administrative support and direction, as also stated by Beltman et al. (2011). Therefore, they suggested that the teacher education program provide more information to teacher candidates. The findings of this study agree with related studies that emphasize the importance of tailoring assessment training to meet the specific needs of real classrooms (Coombe et al., 2012; Gotch & McLean, 2019; Juanjuan & Mohd Yusoff, 2022; Livingston & Hutchinson, 2017; Stiggins & Conklin, 1988; Tsagari, 2016). Coombe et al. (2020) pointed out that teacher assessment training should be extended and sustained to actively involve teachers in deep learning about assessment, potentially enhancing their understanding and practices of assessment. Moreover, they stated that it is important to consider the knowledge base and practice context when providing assessment training and to establish relationships between them. This means that assessment literacy should be developed by considering different educational environments and the specific needs of those environments. Similarly, Stiggins and Conklin (1988) have stated that teacher education programs can prepare teachers for authentic assessment environments when they receive training that aligns with the actual needs of classrooms.

5.3.2 How: Method to Improve Assessment Literacy

Some of the suggestions mentioned by the participants regarding the teacher education program's provision of more quality education in terms of assessment literacy were related to how it should be done, in other words, the methods involved. The suggestions offered by the participants have been categorized into three categories, similar to the study by Juanjuan and Mohd Yusoff (2022): building a collaborative learning community, advocating active engagement, and providing more opportunities regarding assessment. Detailed information on each suggestion is given in the following sections.

5.3.2.1 Building a Collaborative Learning Community

The participants considered collaborative learning to enhance assessment literacy, where ideas are shared, thoughts are developed, and perspectives are broadened through the exchange of experiences. All participants emphasized the importance of collaborating with experienced teachers including mentor teachers. For example, they mentioned the benefits of further collaboration with their mentor teachers, who guided them during their teaching experience in assessment. They stated that observing, discussing, and getting feedback for their assessment methods would help improve their assessment knowledge and practices. From this suggestion of the participants, it can be concluded that engaging in more collaboration with an experienced teacher actively teaching in a real classroom environment can provide pre-service teachers with greater opportunities to enhance their understanding of assessment. These findings align with Sakash and Rodriguez-Brown (2010)'s study. They stated that the meeting of experienced teachers and pre-service teachers affects pre-service teachers' learning positively. Similarly, Du Plessis et al. (2010) emphasized the importance of placing teacher candidates in schools that provide supportive learning environments. Additionally, mentoring educators should acquire suitable practicum, and there needs to be greater transparency regarding the individuals involved, the specific assessment criteria, and the methodologies employed in evaluating students during their teaching practice.

Another collaboration suggestion emphasized by the participants was related to university faculty members. They mentioned that university faculty members also assess teacher candidates, meaning that the knowledge they learn in assessment courses is applied to them. The participants believed it would benefit university faculty members to share their experiences and methods regarding the assessments they have conducted with teacher candidates. Similarly, Allen and Flippo (2002) remarked that university faculty members should demonstrate various assessment strategies in their courses to help pre-service students develop a more comprehensive understanding of the effectiveness of diverse assessment procedure. Considering this suggestion of the participants, being knowledgeable about the assessment process that they have experienced may assist them in reflecting on and making inferences about assessment more easily. In addition, the participants mentioned that it would be better if university faculty members gave them more feedback about the assessment part of assignments and lesson plans. While the study does not provide a direct conclusion, this suggestion could indicate the need for university faculty members to place more emphasis on assessment in teacher education. The results of Ferguson (2011) study showed that providing continuous and detailed feedback to teacher candidates significantly improves their learning.

Another cooperation proposal offered by the participants was cooperation with peers. Participants suggested that more group assignments and discussions be carried out on assessment. As Jackson and Bruegmann (2009) stated, peer collaboration enhances learning by increasing motivation and facilitating the exchange of ideas among peers. The literature and the participants' suggestions indicate that incorporating more group work and discussions among pre-service teachers in assessment training would contribute to developing their assessment knowledge and skills.

The last suggestion from the participants' regarding collaboration was to engage with evaluation experts. For instance, they mentioned that collaborating with experts from the Ministry of National Education would allow teacher candidates to benefit from the experience and expertise of assessment professionals. These findings support the idea that cooperation between teacher candidates and experts can contribute to their assessment development. In the relevant literature, it has been found that collaborative learning increases assessment knowledge and skills (Alkharusi et al., 2011; Baker &

Riches, 2018; Lukin et al., 2004; Nolen et al., 2011; Prizovskaya, 2017; Xu, 2019; Yamtim & Wongwanich, 2014).

For instance, Baker and Riches (2018) conducted a study during a series of language assessment workshops with the goal of investigating the development of language assessment literacy among teachers and language assessors in this collaborative setting. Based on their findings, they concluded that this collaborative environment contributed to the development of participants' assessment literacy.

5.3.2.2 Advocating Active Engagement

As Mattsson et al. (2011) noted, translating theoretical knowledge into practice would allow teacher candidates to experience real classroom situations, thus contributing to their professional development and making them more competent. Bullough Jr (2002) expressed the importance of practice as follows, "theory could not be separated from practice; there is no escape" (p. 18). Therefore, it is vital to practice as well as theoretical knowledge. Mertler (2009) implemented a two-week practice-centered assessment workshop for in-service teachers and examined its impact on teachers' assessment literacy. The study's findings indicated that the teachers' assessment literacy increased in the post-workshop test compared to their pre-workshop level. The results of the current study also demonstrated the need for more practice or active engagement, focused explicitly on assessment in teacher education programs to enhance the quality of assessment education provided. To this end, all participants suggested encouraging teacher candidates to actively participate in the assessment processes by taking on more responsibilities during their internship. Considering this suggestion of the participants, the internship can be a valuable opportunity to put assessment knowledge into practice. The internship provides pre-service teachers with teaching experience, during which they can develop their assessment literacy along with other teaching skills. Aligning with the participants' suggestions, Volante and Fazio (2007) also argued for the importance of gaining assessment experience in practical teaching environments. Therefore, it can be beneficial to involve teacher candidates more in assessment during the internship program offered in teacher education programs.

Another suggestion from many participants was the need for a more student-centered assessment course. They pointed out that teacher candidates have limited participation in the class, as most of the content is delivered by instructors in the evaluation course. They emphasized that with a student-centered assessment course, pre-service teachers would be more engaged in their own learning processes, leading to more meaningful and long-lasting learning outcomes. Therefore, they recommended the assessment courses be designed to promote active participation and meaningful engagement in class discussions for teacher candidates. In their study, Boyd and Donnarumma (2018) implemented a student-centered assessment lesson model, where teachers were encouraged to formulate choices independently and actively make assessments. The study's findings indicated that this approach enhanced teachers' confidence in assessment, thus significantly improving assessment literacy. In line with the literature, student-centered assessment courses may greatly contribute to assessment literacy development.

Another suggestion from the participants was to ensure that teacher candidates attend parent meetings. They mentioned that the teacher education program was lacking in communicating with parents about assessment. For this reason, the participants, especially those who are teachers, stated that the teacher education program was insufficient in preparing them to communicate with parents in real life. They suggested that pre-service teachers could at least attend parent meetings as observers during the internship course. The participation of teacher candidates in parent meetings during their teaching experience can help them observe the communication between an experienced teacher and parents, as well as assist parents in recognizing their expectations regarding assessment from the teacher. Therefore, they can be more self-confident and conscious in communicating with parents about the evaluation results when they start their professional life.

Another suggestion offered by the participants was for them to have back-to-back teaching experience. The participants highlighted that during their undergraduate education, they often focused on planning lessons for individual class sessions without considering the assessment results for the subsequent lessons. They recognized this as a significant deficiency and recommended that the teacher education program should provide opportunities for them to plan and teach

subsequent lessons, considering the assessment results. Considering this suggestion, it can be argued that it would be beneficial for teacher education programs to provide teacher candidates with the opportunity to develop and implement lesson plans for a longer duration. In this process, incorporating active assessment practices and designing lesson plans based on assessment results can enhance a more meaningful understanding of assessment information.

The last suggestion of the participants on this subject was to implement all assessment processes to the real student(s) in a way that is linked like a long-term project. The participants emphasized during the interviews that each sub-dimension (six standards) asked by the researcher was interconnected and that transitioning from one to another should be done appropriately without skipping any. Therefore, they believed that implementing all of these processes by teacher candidates would significantly enhance their assessment literacy levels. In this regard, it can be said that preparing a long-term project in undergraduate education will be effective.

The findings of the current study align with the existing literature. Numerous studies indicate that engaging in practical application, which involves putting theoretical knowledge into practice, enhances assessment literacy and fosters meaningful learning (Alkharusi et al., 2011; Beziat & Coleman, 2015; Brookhart, 1999; Dehqan & Asadian Sorkhi, 2020; DeLuca & Klinger, 2010; Graham, 2005; Maclellan, 2004; Mertler & Campbell, 2005; Ogan-Bekiroglu & Suzuk, 2014; Volante & Fazio, 2007; Xu & Brown, 2016). The findings of the study can support the literature by providing a detailed framework about what kind of experiences that teacher education programs can offer in terms of practicality for assessment literacy.

5.3.2.3 Providing More Learning Opportunities Regarding Assessment

As stated by Malone (2008), learning communities, such as conferences, seminars, panels, and workshops enhance individuals' professional growth by providing access to current advancements in their domains and other sources of knowledge. This enables individuals to improve their professional knowledge, skills, and competencies. The findings of this study also demonstrated that providing these opportunities to teachers and teacher candidates would help them acccess higher-quality education in

terms of assessment literacy. For this reason, the participants suggested that their universities should offer teacher candidates such opportunities to improve themselves in assessment. Considering this suggestion of the participants, participating in such learning environments can help pre-service teachers stay updated with the latest developments in the assessment, expand their professional networks with assessment experts, gain new ideas, and further their professional development. These findings are consistent with the results of other studies that recommend providing professional development opportunities to support assessment literacy development (Akayuure, 2021; Elshawa et al., 2017; Klinger et al., 2012; Lukin et al., 2004; Mertler, 2009; Tsagari, 2016). In addition, the participants suggest engaging in independent research and readings on assessment for their professional development in assessment literacy. In light of this, it can be said that teacher education programs offering independent research and reading resources on assessment to teacher candidates can contribute to their development of assessment literacy. teacher education programs could establish a library dedicated explicitly to assessment.

5.3.3 Why: Orientation to Improve Assessment Literacy

One suggestion put forward by the participants relates to the orientation of teacher education programs. The theme of "why (orientation)" emphasizes the importance of assisting educators in reevaluating their role as assessors and supporting them in becoming reflective practitioners, as noted by Juanjuan and Mohd Yusoff (2022). However, in the current study, there was only one suggestion in this category put forward by the participants, and that is "cultivating reflective practitioners." Detailed information is given in the following section.

5.3.3.1 Cultivating Reflective Practitioner

The reflection supports teachers and teacher candidates in enhancing their professional understanding and skills in assessment (Klinger et al., 2012) while reflecting on assessment practices empowers them to cultivate intuitive principles and guidelines in assessment (Yan et al., 2018). In their research, Ogan-Bekiroglu

and Suzuk (2014) noted that pre-service teachers might not be aware of the significance of assessment for education. In parallel with the findings of the literature, the participants have argued that teacher education programs need to place more emphasis on the importance of assessment. This suggestion by the participants would indicate the necessity of teacher education programs to convey the importance of assessment better to transform teacher candidates into reflective practitioners focused on assessment.

5.4 Conclusions

Both the quantitative and qualitative findings of the study indicate deficiencies in the teacher education program where the study was conducted in terms of assessment literacy education. The lack of a significant difference between the assessment literacy scores of both pre-service and in-service middle school mathematics teachers and the insights gathered through interviews with the participants point to both theoretical and practical shortcomings.

Moreover, the study's results might suggest that both pre-service and in-service teachers share similar deficiencies or limitations in their grasp of assessment concepts, techniques, and practices. This implies that the teacher education program for both groups may not effectively address the cultivation of assessment literacy. Consequently, it could signal a requirement for a comprehensive review and enhancement of the teacher education curriculum by university faculty members to better equip teachers with essential assessment knowledge and skills.

Furthermore, the study's outcomes may highlight a potential misalignment between the content covered in teacher education programs and the practical needs of educators in real classroom settings. This, in turn, could underscore the necessity for a closer alignment between the program curriculum and the challenges that teachers face concerning assessment in their day-to-day teaching experiences.

5.5 Implications for Educational Practices

Researchers and teacher educators can learn from the results of the current study about the shortcomings of the teacher preparation program in developing pre-service and in-service middle school mathematics teachers. Therefore, this section will present some implications for the teacher education programs, university faculty members, and the Ministry of National Education (MoNE) regarding the development of assessment literacy.

The findings of this study demonstrated that the assessment training offered in teacher education has both theoretical and practical deficiencies. According to pre-service and in-service teachers, the education they received for assessment in undergraduate education was superficial and insufficient. They emphasized that assessment is a complex subject with numerous sub-topics, thus indicating that a single course in assessment during undergraduate education is inadequate. Additionally, they argued that the absence of elective courses on assessment in their programs is a significant limitation for teacher candidates who seek to enhance their expertise in this domain. To address these issues, teacher education programs can increase the number of compulsory and elective courses offered for assessment. Furthermore, the assessment course curriculum can be revised, taking into account the theoretical deficiencies highlighted by pre-service and in-service teachers.

Besides the assessment course, another point raised by pre-service and in-service teachers was the neglect of the assessment dimension in other courses related to assessment. They stated that university faculty members did not establish connections with assessment in courses other than the specific assessment course. They also mentioned that some assignments and lesson plan included assessments, but university faculty members did not provide any feedback on the assessment part of these tasks. Lesson plans that take into account all stages of a course, including assessment, are essential tasks that enable the development of teacher candidates in using evaluation to improve instruction. For this reason, it is crucial to develop assessment literacy that university lecturers review and provide feedback on the lesson plans, considering the teaching method and the assessment aspect. Therefore, university faculty members can help teacher candidates establish connections

between teaching and assessment by emphasizing assessment in courses other than the specific assessment course. Additionally, they can support the development of teacher candidates by providing feedback on the assessment dimension of the assignments or lesson plans they create.

The findings of the current study showed that the assessment training offered by the teacher education program also has shortcomings in practice. Pre-service and in-service teachers stated that the teacher education program does not provide an environment and opportunities to apply the theoretical knowledge they have learned. Not using theoretical knowledge caused them to be unable to establish a connection between theory and practice and hindered their absorption of theoretical knowledge. To overcome this deficiency, it is necessary to offer teacher training to teacher candidates where they can practice more. In undergraduate education, assessment-related assignments may increase when preparing a lesson plan or tasks contributing to the overall grade. Pre-service teachers may be expected to perform project-based tasks such as developing exams, tests, and rubrics, grading and analyzing exams, in which they will produce assessment materials/tools/practices that are valid, fair, and useful for students.

Additionally, based on the results of the present study, there is a lack of practice in assessment in the internship experience, where teacher candidates have the opportunity to apply their theoretical knowledge. Pre-service and in-service teachers stated that while they needed to gain teaching experience during their internship, they had very few opportunities to do so, and they mostly passed their internship by observing. When they obtained this experience, they also mentioned that they had very few opportunities for assessment-related practices, and therefore, the internship did not contribute significantly to their assessment literacy. However, just like the in-service teachers who participated in the study pointed out, teacher candidates must practice with real students when they become teachers to conduct appropriate and high-quality assessments. Therefore, in the internship course, teacher candidates should be provided with more opportunities to feel like real teachers and engage in more assessment practices. Due to the heavy workload of the middle school curriculum in Turkey, it may be challenging to offer teacher candidates more teaching experiences in an actual classroom setting. In this regard, a collaboration between the

MoNE and universities providing education on assessment can be established. A plan and framework can be developed to make this possible, followed by implementing an official policy.

Moreover, the current study put forward that the mentor teachers who guide the pre- service teachers during the internship experience do not adequately contribute to the development of their assessment literacy. Participants pointed out that mentor teachers do not use many assessment methods in their lessons. As a result, they cannot achieve significant learning about assessment by observing the experienced mentor teachers. They also indicated that they did not establish communication and cooperation with their mentors on issues such as analyzing, interpreting, scoring, and discussing assessment results outside of the classroom, apart from the assessment they conducted for the students in the classroom. Considering all these findings, it can be suggested that mentor teachers should share their experiences and knowledge about assessment with teacher candidates more. For instance, mentor teachers can ask teacher candidates to score the assessments they have conducted on students. Then, by comparing their scoring with the teacher candidates, they can share their experiences on the methods or aspects they paid attention to. This collaboration between mentor teachers and teacher candidates can provide the candidates with more practice opportunities and enable them to benefit from the knowledge of an experienced teacher, making them more competent in assessment.

In addition, teacher education programs can collaborate with the MoNE to offer assessment-focused seminars, workshops, panels, or conferences where teacher candidates can further develop themselves. These training sessions can involve assessment experts at the MoNE and experienced teachers who can share their experiences and guide teacher candidates.

Lastly, there are implications for the recommendations provided by the participants to address the shortcomings and improve the quality of assessment literacy in the teacher education program. By considering the deficiencies and suggestions mentioned by pre-service and in-service teachers in the overall context and specific sub-dimensions, teacher education programs, and university faculty members, the MoNE and researchers can revise the assessment education in teacher training.

5.6 Recommendations for Further Studies

Additionally, the current study offers suggestions for future research based on its findings. To begin with, studies involving a greater number of pre-service and in-service teachers from different teacher education programs can be conducted. As the number of participants in the research increases, more diverse and varied data will be obtained, contributing to the enhancement of assessment literacy education provided in undergraduate programs. Furthermore, the pre-service and in-service teachers in the current study are participants who have either studied or graduated from the same university. To obtain richer data for the development of teacher education programs, an investigation can be conducted involving pre-service and in-service teachers from different universities.

Secondly, the design of the study can be altered to allow for future research opportunities. Although the current study includes quantitative analyses, it is primarily qualitative. With more participants, a mixed-method approach involving both qualitative and quantitative methods can be employed. In more detailed quantitative research, a specific examination can be conducted on the relationship between competency areas. Alternately, experimental research methods can be employed to provide assessment literacy training to a specific group, considering competency areas, and differences between control and treatment groups can be examined, or pre-tests and post-tests can be administered. Moreover, the assessment literacy of pre-service and in-service teachers can be compared quantitavely. This way, richer data can be obtained by analyzing quantitative and qualitative data.

Thirdly, the impact of other factors, such as participants' educational background, the type of school they work in, their average grades, gender, beliefs, attitudes, or self-efficacy towards assessment, can be investigated in this context on their assessment literacy. Thus, it can be determined whether different factors influence the assessment literacy of pre-service and in-service teachers.

Fourthly, a similar study could be conducted using a data collection tool that measures standard 7, which is "teachers need to possess the competence to identify assessment information that is unethical, illegal, or otherwise inappropriate," as found in the

Assessment Literacy Inventory developed by Mertler and Campbell (2005). Because the Turkish-translated assessment literacy inventory used in the current study does not include questions related to the 7th standard. Consequently, data regarding pre-service and in-service middle school mathematics teachers' assessment literacy levels, deficiencies in the teacher training program, and recommendations related to this standard could not be obtained. This would also allow the teacher education program to be assessed and improved in terms of competency area 7.

Fifthly, according to the results of the current study, both middle school mathematics pre-service and in-service teachers obtained their lowest scores on the assessment literacy inventory in the fifth competency area, which is "creating accurate techniques for grading students based on their assessments." A study solely focused on this competency area, aiming to understand the reasons behind the low scores, can be conducted.

Finally, a similar study could be conducted by changing the type of participants. Besides pre-service and in-service teachers, teacher' educators assessment literacy could also be included in the study to gather their thoughts and insights on the shortcomings and recommendations regarding assessment literacy education offered in teacher education programs.

5.7 Limitations of the Study

There are some limitations of the study. Firstly, although this study's quantitative data collection tool was applied to 34 pre-service and 41 in-service teachers, interviews were conducted with 5 pre-service and 5 in-service teachers. As a result, the researcher could not interview all participants who participated in the task implementation, and the generalizability of the current study's conclusions is limited.

Secondly, the participants of the study, especially in-service teachers, sometimes had difficulty remembering the courses they took for assessment in undergraduate education. Although this situation limited the findings presented in the study, it was also essential data for the study. Although the participants stated that they remembered very clearly some of the knowledge and experiences they gained

in undergraduate education, they considered some issues/situations that they had difficulty remembering as the lack of undergraduate education regarding that issue. Accordingly, in future studies, a study on assessment literacy with in-service teachers could be conducted to investigate whether in-service teachers vividly recall certain information they learned during their undergraduate education while forgetting other information.

Thirdly, the academic language of the university that the participants graduated from is English. However, both the assessment literacy inventory and the interview questions were prepared in Turkish, their mother tongue. The reason for this is the belief that participants, despite receiving education in English at the university, may forget English over time and would better comprehend questions prepared in their native language, Turkish. Since they had a better command of the English words and terms they learned in their undergraduate education, they were likely to have difficulty finding the Turkish equivalent of the words and terms. In order to overcome this limitation, the researcher added the English equivalents in parentheses next to the words and terms related to assessment in the inventory. Furthermore, during the interviews, the researcher tried to provide necessary explanations when participants encountered difficulties in comprehending words and terms.

Finally, despite being participants in the same university, pre-service and in-service middle school mathematics teachers have been part of the university's teacher education program in different years. Based on the academic catalogs of the university's courses offered in various years, which indicated that the course contents for different years had similar learning objectives and content, the researcher assumed that both groups had similar undergraduate education experiences. This assumption is also one of the limitations of the study.

REFERENCES

AFT, NCME, & NEA. (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice*, *9*(4), 30–37.

Akayuure, P. (2021). Classroom assessment literacy levels of mathematics teachers in ghanaian senior high schools. *Contemporary Mathematics and Science Education*, *2*(2), 1–9.

Alkharusi, H., Kazem, A. M., & Al-Musawai, A. (2011). Knowledge, skills, and attitudes of preservice and inservice teachers in educational measurement. *Asia-Pacific Journal of Teacher Education*, *39*(2), 113–123.

Allen, D. D., & Flippo, R. F. (2002). Alternative assessment in the preparation of literacy educators: Responses from students. *Reading Psychology*, *23*(1), 15–26.

Altbach, P. G., Reisberg, L., & Rumbley, L. E. (2019). *Trends in global higher education: Tracking an academic revolution* (Vol. 22). Brill.

Anderson, L. W. (2003). *Classroom assessment: Enhancing the quality of teacher decision making*. Routledge.

Andrade, H. G. (1997). Understanding rubrics. *Educational Leadership*, *54*(4), 14–17.

Archer, E. (2017). The assessment purpose triangle: Balancing the purposes of educational assessment. *Frontiers in Education*, *2*(41), 1–7.

Arter, J. A. (2001). *Washington assessment professional development program evaluation results. (*ED456124*)*. ERIC. https://files.eric.ed.gov/fulltext/ED456124.pdf

Atılgan, H., Kan, A., & Doğan, N. (2007). *Eğitimde ölçme ve değerlendirme*. Anı Yayıncılık.

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, *27*(1), 10–20.

Ayalon, M., & Wilkie, K. J. (2020). Developing assessment literacy through approximations of practice: Exploring secondary mathematics pre-service teachers developing criteria for a rich quadratics task. *Teaching and Teacher Education*, *89*, 1–14. https://doi.org/10.1016/j.tate.2019.103011

Ayvaz-Tuncel, Z., & Çobanoğlu, F. (2018). In-service teacher training: Problems of the teachers as learners. *International Journal of Instruction*, *11*(4), 159–174.

Babchuk, W. A. (1996). Glaser or Strauss? Grounded theory and adult education. *Proceedings of the 15th Annual Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education*, *12*(2), 1–6.

Baird, J.-A., Andrich, D., Hopfenbeck, T. N., & Stobart, G. (2017). Assessment and learning: Fields apart? *Assessment in Education: Principles, policy & practice*, *24*(3), 317–350.

Baker, B. A., & Riches, C. (2018). The development of EFL examinations in haiti: Collaboration and language assessment literacy development. *Language Testing*, *35*(4), 557–581.

Barnes, N., Fives, H., & Dacey, C. M. (2017). US teachers' conceptions of the purposes of assessment. *Teaching and Teacher Education*, *65*, 107–116.

Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. ÖSYM Yayınları.

Beck, M. D. (1980). Student, teacher, and measurement specialist attitudes toward standardized tests: A summary of three surveys. In L. M. Rudner (Ed.), *Testing in our schools* (pp. 1–13). National Institute of Education.

Beltman, S., Mansfield, C., & Price, A. (2011). Thriving not just surviving: A review of research on teacher resilience. *Educational Research Review*, *6*(3), 185–207.

Beziat, T. L., & Coleman, B. K. (2015). Classroom assessment literacy: Evaluating pre-service teachers. *The Researcher*, *27*(1), 25–30.

Biggs, J. (1998). Assessment and classroom learning: A role for summative assessment? *Assessment in Education: Principles, Policy & Practice*, *5*(1), 103–110.

Bijsterbosch, E., Béneker, T., Kuiper, W., & van der Schee, J. (2019). Teacher professional growth on assessment literacy: A case study of prevocational

geography education in the Netherlands. *The Teacher Educator*, *54*(4), 420–445.

Black, P. (2002). *Testing: Friend or foe?: Theory and practice of assessment and testing*. Routledge.

Black, P., Harrison, C., & Lee, C. (2003). *Assessment for learning: Putting it into practice*. McGraw-Hill.

Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. Granada Learning.

Blaich, C., & Wise, K. (2011). *From gathering to using assessment results: Lessons from the Wabash National Study* (NILOA Occasional Paper No. 8). University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment. Urbana, IL.

Blömeke, S., & Delaney, S. (2012). Assessment of teacher knowledge across countries: A review of the state of research. *ZDM-Mathematics Education, 44*(3), 223–247.

Bloomfield, J., & Fisher, M. J. (2019). Quantitative research design. *Journal of the Australasian Rehabilitation Nurses Association*, *22*(2), 27–30.

Boikhutso, K. (2010). The theory into practice dilemma: Lesson planning challenges facing botswana student-teachers. *Improving Schools*, *13*(3), 205–220.

Bol, L., & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education*, *80*(2), 145–163.

Boud, D. (2015). Feedback: Ensuring it leads to enhanced learning. *The Clinical Teacher*, *12*, 3–7. https://doi.org/10.1111/tct.12345

Boyd, E., & Donnarumma, D. (2018). Assessment literacy for teachers: A pilot study investigating the challenges, benefits and impact of assessment literacy training. In *Teacher involvement in high-stakes language testing* (pp. 105–126). Springer. https://doi.org/10.1007/978-3-319-77177-9_7

Bozkurt, E. (1995). Eğitimde değerlendirmenin gerekliliği. *Kuram ve Uygulamada Eğitim Yönetimi*, *4*(4), 531–534.

Brookhart, S. M. (1999). Teaching about communicating assessment results and grading. *Educational Measurement: Issues and Practice*, *18*(1), 5–13.

Brookhart, S. M. (2001). *The standards and classroom assessment research. (*ED451189*)*. ERIC. https://files.eric.ed.gov/fulltext/ED451189.pdf

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, *30*(1), 3–12.

Brown, G. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. Nova Science Publishers.

Brown, G. A., Bull, J., & Pendlebury, M. (2013). *Assessing student learning in higher education*. Routledge.

Brown, S. (2014). Assessment literacy: The foundation for improving student learning. *Assessment & Evaluation in Higher Education*, *39*(3), 395–396. https://doi.org/10.1080/02602938.2013.820564

Bullough Jr, R. V. (2002). Practicing theory and theorizing practice in teacher education. In *Teaching about teaching* (pp. 27–45). Routledge.

Burger, D. (2000). *Assessment and accountability. PREL briefing paper. (*ED466492*)*. ERIC. https://files.eric.ed.gov/fulltext/ED466492.pdf

Bütüner, S. Ö., Uzun, S., & Yiğit, N. (2010). A comparison of the results of TIMSS 1999-2007: The most successful five countries-turkey sample. *Elementary Education Online*, *9*(3), 1174–1188.

Büyüköztürk, Ş., Kılıç-Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2013). *Bilimsel araştırma yöntemleri*. Pegem Yayıncılık.

Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, *37*(2), 99–114.

Çambay, Ö. (2016). Hizmet içi ve hizmet öncesi öğretmenlerinin değerlendirme okuryazarlığı ve değerlendirme okuryazarlığına ilişkin öz-yeterliklerinin araştırılması. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, *13*(1), 16–40.

Campbell, C., Murphy, J., & Holt, J. (2002). Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers. *Annual meeting of the mid-western educational research association, Columbus, OH*.

Campbell, C., & Evans, J. A. (2000). Investigation of preservice teachers' classroom assessment practices during student teaching. *The Journal of Educational Research*, *93*(6), 350–355.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in higher education*, *36*(4), 395–407.

Çelik, F., & Önal, A. S. (2005). Öğretimde planlama ve değerlendirme dersi öğretim programının değerlendirilmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, *18*(18), 39–59.

Çelikkaya, T., Karakuş, U., & Demirbaş, Ç. (2010). Sosyal bilgiler öğretmenlerinin ölçme-değerlendirme araçlarını kullanma düzeyleri ve karşılaştıkları sorunlar. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, *11*(1), 57–76.

Chand, R., Alasa, V. M., Chitiyo, J., & Pietrantoni, Z. (2022). Preparation of pre-service teachers: Assessment of generation z students. In J. Keengwe (Ed.), *Handbook of research on digital-based assessment and innovative practices in education* (pp. 116–130). IGI Global.

Chapelle, C. A., Kremmel, B., & Brindley, G. (2010). Assessment. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 247–266). Routledge.

Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership*, *60*(1), 40–44.

Charmaz, K. (2008). *Qualitative psychology: A practical guide to research methods* (2nd ed.). Sage.

Chase, C. I. (1999). *Contemporary assessment for educators*. Addison-Wesley Longman.

Cizek, G. J. (2010). An introduction to formative assessment: History, characteristics, and challenges. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 3–17). Routledge.

Clark, A. K., Nash, B., & Karvonen, M. (2022). Teacher assessment literacy: Implications for diagnostic assessment systems. *Applied Measurement in Education*, *35*(1), 17–32.

Coombe, C., Troudi, S., & Al-Hamly, M. (2012). Foreign and second language teacher assessment literacy: Issues, challenges, and recommendations. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoynoff (Eds.), *The Cambridge guide to second language assessment* (pp. 20–29). Cambridge University Press.

Coombe, C., Vafadar, H., & Mohebbi, H. (2020). Language assessment literacy: What do we need to learn, unlearn, and relearn? *Language Testing in Asia*, *10*(3), 1–16.

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Sage.

CSAI. (2018). *Valid and reliable assessments. (*ED588476*)*. ERIC. https://files.eric.ed.gov/fulltext/ED588476.pdf

Darr, C. (2005). A hitchhiker's guide to reliability. *Set: Research Information for Teachers*, (3), 59–60. https://doi.org/10.18296/set.0623

Dehqan, M., & Asadian Sorkhi, S. R. (2020). Pre-service and in-service teachers' knowledge and practice of assessment literacy: A dweller in an ivory tower. *Issues in Language Teaching*, *9*(2), 347–375.

DeLuca, C., & Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education curriculum. *Journal of Teacher Education*, *64*(4), 356–372.

DeLuca, C., Chavez, T., Bellara, A., & Cao, C. (2013). Pedagogies for preservice assessment education: Supporting teacher candidates' assessment literacy development. *The Teacher Educator*, *48*(2), 128–142.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 419–438.

DeLuca, C., Lapointe-Mcewan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, *28*, 251–272.

DeLuca, C., & McEwen, L. (2007). Evaluating assessment curriculum in teacher education programs: An evaluation process paper. *Paper presented at the annual Edward F. Kelly Evaluation Conference, April, in Ottawa, ON*.

Demirtaşlı, R. N. (Ed.). (2014). *Eğitimde ölçme ve değerlendirme*. Edge Akademi.

Deneen, C., & Deneen, J. R. (2008). *Assessing student achievement: A guide for teachers and administrators*. R & L Education.

Deneen, C. C., & Brown, G. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, *3*(1), 1–14. https://doi.org/10.1080/2331186X.2016.1225380

Deneen, C. C., & Hoo, H.-T. (2023). Connecting teacher and student assessment literacy with self-evaluation and peer feedback. *Assessment & Evaluation in Higher Education*, *48*(2), 214–226.

DePaul, A. (2000). *Survival guide for new teachers: How new teachers can work effectively with veteran teachers, parents, principals, and teacher educators*. US Department of Education, Office of Educational Research; Improvement.

Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical education*, *37*(9), 830–837.

Du Plessis, E. C., Marais, P., Van Schalkwyk, A., & Weeks, F. (2010). Adapt or die: The views of Unisa student teachers on teaching practice at schools. *Africa Education Review*, *7*(2), 323–341.

Ediger, M. (2000). Measurement/evaluation courses in teacher education. *Education*, *121*(1), 169–169.

Eğri, G. (2006). *Coğrafya öğretmenlerinin ölçme değerlendirme yapabilme yeterliliği* (Master's thesis). Gazi University.

Elshawa, N. R., Abdullah, A. N., & Rashid, S. M. (2017). Malaysian instructors' assessment beliefs in tertiary ESL classrooms. *International Journal of Education and Literacy Studies*, *5*(2), 29–46.

Elturki, E. (2020). A systematic process for assessing assessment. *English Teaching Forum*, *58*(4), 12–21.

Emaliana, I. (2017). Teacher-centered or student-centered learning approach to promote learning? *Jurnal Sosial Humaniora*, *10*(2), 59–70.

Eret, E. (2013). *An assessment of pre-service teacher education in terms of preparing teacher candidates for teaching* (Doctoral dissertation). Middle East Technical University.

Ergül, A. Ö., & Çetin, S. (2021). Measurement and assessment literacy levels of teachers in the context of certain factors. *Adıyaman University Journal of Educational Sciences*, *11*(1), 26–35.

Essomba, M. A. (2010). Teacher education for diversity in Spain: Moving from theory to practice. In T. Burns, V. Shadoian-Gersing, J. Bouch, T. Walsh, & C. Davis (Eds.), *Educating teachers for diversity* (pp. 219–236). OECD Publishing. https://doi.org/10.1787/9789264079731-13-en

Estes, C. A. (2004). Promoting student-centered learning in experiential education. *Journal of experiential education*, *27*(2), 141–160.

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education*, *36*(1), 51–62.

Flórez, M. T., & Sammons, P. (2013). *Assessment for learning: Effects and impact (*ED546817*)*. ERIC. https://files.eric.ed.gov/fulltext/ED546817.pdf

Fraenkel, J. R., Wallen, N. E., Hyun, H. H., et al. (2011). *How to design and evaluate research in education* (Vol. 7). McGraw-Hill.

Gardner, J. (2012). *Assessment and learning* (2nd ed.). Sage.

Gay, L. R., Mills, G. E., & Airasian, P. W. (2006). *Educational research: Competencies for analysis and applications* (8th ed.). Pearson Merrill Prentice Hall.

Gee, J. (2005). The new literacy studies: From 'socially situated' to the work of the social. In R. Ivanic (Ed.), *Situated literacies: Theorising reading and writing in context* (pp. 177–193). Routledge.

Gotch, C. M., & McLean, C. (2019). Teacher outcomes from a statewide initiative to build assessment literacy. *Studies in Educational Evaluation*, *62*, 30–36.

Gottheiner, D. M., & Siegel, M. A. (2012). Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education*, *23*, 531–557.

Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through preservice teacher education. *Teaching and Teacher Education*, *21*(6), 607–621.

Gravetter, F. J., & Wallnau, L. B. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Wadsworth Cengage Learning.

Gregory, I. (2003). *Ethics in research*. Continuum.

Grossman, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal, 45*(1), 184–205.

Gu, P. Y. (2014). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a foreign language in a Chinese secondary school? *Assessment in Education: Principles, Policy & Practice*, *21*(3), 286–305.

Gül, E. (2011). *İlköğretim öğretmen adaylarının ölçme-değerlendirme okuryazarlığı ve ölçme-değerlendirmeye ilişkin tutumlarının belirlenmesi* (Master's thesis). Fırat University.

Gullickson, A. R. (2003). *The student evaluation standards: How to improve evaluations of students*. Corwin Press.

Gürsoy, G. (2017). Ölçme ve değerlendirme okuryazarlığı: Kavramsal bir analiz. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, *6*(1), 281–316.

Guskey, T. R., & Bailey, J. M. (2001). *Developing grading and reporting systems for student learning*. Corwin Press.

Guskey, T. (2003). How classroom assessments improve learning. *Educational Leadership*, *60*(5), 6–11.

Haladyna, T. M., & Downing, S. M. (Eds.). (2006). *Handbook of test development* (1st ed.). Routledge. https://doi.org/10.4324/9780203874776

Harrison, R., Meyer, L., Rawstorne, P., Razee, H., Chitkara, U., Mears, S., & Balasooriya, C. (2022). Evaluating and enhancing quality in higher education teaching practice: A meta-review. *Studies in Higher Education*, *47*(1), 80–96.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, *89*(2), 140–145.

Hill, M. (1999). Assessment in self-managing schools: Primary teachers balancing learning and accountability demands in the 1990s. *New Zealand Journal of Educational Studies*, *34*(1), 176–185.

Hogan, T. P. (2007). *Educational assessment: A practical introduction*. John Wiley.

Huang, J., & He, Z. (2016). Exploring assessment literacy. *Higher Education of Social Science*, *11*(2), 18–27.

Husebo, D. (2012). Bridging theory and practice in norwegian teacher education through action research. *Educational Action Research*, *20*(3), 455–471.

Irons, A., & Elkington, S. (2021). *Enhancing learning through formative assessment and feedback*. Routledge.

Jackson, C. K., & Bruegmann, E. (2009). Teaching students and teaching each other: The importance of peer learning for teachers. *American Economic Journal: Applied Economics*, *1*(4), 85–108.

Janssen, F., Grossman, P., & Westbroek, H. (2015). Facilitating decomposition and recomposition in practice-based teacher education: The power of modularity. *Teaching and Teacher Education*, *51*, 137–146.

Joint Committee for Standards in Educational Evaluation. (2003). *Student evaluation standards*. Corwin Press.

Juanjuan, G., & Mohd Yusoff, N. (2022). The shared features of effective improvement programmes for teachers' assessment literacy. *Pedagogies: An International Journal*. https://doi.org/10.1080/1554480X.2022. 2084398

Kaplan, G. (2015). *An investigation of preservice middle school mathematics teachers' formative assessment approaches through lesson planning* (Doctoral dissertation). Middle East Technical University.

Karaman, P. (2014). *Öğretmen adaylarının ölçme-değerlendirme okuryazarlıklarının belirlenmesi ve mikro-öğretim yoluyla geliştirilmesi* (Doctoral dissertation). Çanakkale Onsekiz Mart University.

Karaman, P., & Şahin, Ç. (2014). Öğretmen adaylarının ölçme-değerlendirme okuryazarlıklarının belirlenmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, *15*(2), 175–189.

Kehoe, J. (1994). Basic item analysis for multiple-choice tests. *Practical Assessment, Research, and Evaluation*, *4*(10), 1–3. https://doi.org/10.7275/07zg-h235

Khadijeh, B., & Amir, R. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education*, *3*(1), 139–146.

Kim, T. K. (2015). T test as a parametric statistic. *Korean Journal of Anesthesiology*, *68*(6), 540–546.

Klinger, D. A., DeLuca, C., Luu, K., & Sun, Y. (2012). Assessment for learning in the classroom: Barriers to implementation and possibilities for teacher professional learning. *Assessment Matters*, *4*, 5–29.

Koh, K., & Velayutham, R. L. (2009). Improving teachers' assessment literacy in singapore schools: An analysis of teachers' assessment tasks and student work. *Research Brief*, *9*(2), 1–4.

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education*, *22*(3), 255–276. https://doi.org/10.1080/10476210.2011.593164

Korthagen, F. (2010). The relationship between theory and practice in teacher education. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (pp. 669–675). Elsevier.

Krumsvik, R., & Smith, K. (2009). Videopapers–an attempt to narrow the notorious gap between theory and practice in teacher education. *Technology, Pedagogy and Education*, *18*(3), 269–278.

Kuran, K., & Kanatlı, F. (2009). Alternatif ölçme değerlendirme teknikleri konusunda sınıf öğretmenlerinin görüşlerinin değerlendirilmesi/the evaluation of classroom teachers' opinions on the alternative assessments techniques. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *6*(12), 209–234.

Kyttälä, M., Björn, P. M., Rantamäki, M., Lehesvuori, S., Närhi, V., Aro, M., & Lerkkanen, M.-K. (2022). Assessment conceptions of Finnish pre-service teachers. *European Journal of Teacher Education*, 1–19. https://doi.org/10.1080/02619768.2022.2058927

Ladson-Billings, G. (2009). Opportunity to teach: Teacher quality in context. In D. H. Gitomer (Ed.), *Measurement issues and assessment for teaching quality* (pp. 206–222). Los Angeles: Sage.

Lai, E. R., & Waltman, K. (2008). Test preparation: Examining teacher perceptions and practices. *Educational Measurement: Issues and Practice*, *27*(2), 28–45.

Lazer, S. (2009). *Measuring teacher and teaching quality: Considerations and next steps—synthesis of section ii*. Sage.

Lesson planning | Centre for teaching excellence [n.d.]. (n.d.). https://cte.smu.edu.sg/approach-teaching/integrated-design/lesson-planning

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Livingston, K., & Hutchinson, C. (2017). Developing teachers' capacities in assessment through career-long professional learning. *Assessment in Education: Principles, Policy & Practice*, *24*(2), 290–307.

Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy & Practice*, *25*(5), 442–467.

Luft, J. A. (1999). Rubrics: Design and use in science teacher education. *Journal of Science Teacher Education*, *10*(2), 107–121.

Lukin, L. E., Bandalos, D. L., Eckhout, T. J., & Mickelson, K. (2004). Facilitating the development of assessment literacy. *Educational Measurement: Issues and Practice*, *23*(2), 26–32. https://doi.org/10.1111/j.1745-3992.2004.tb00156.x

Macken, S., MacPhail, A., & Calderon, A. (2020). Exploring primary pre-service teachers' use of 'assessment for learning' while teaching primary physical education during school placement. *Physical Education and Sport Pedagogy*, *25*(5), 539–554.

Maclellan, E. (2004). Initial knowledge states about assessment: Novice teachers' conceptualisations. *Teaching and Teacher Education*, *20*(5), 523–535.

Maki, P. L. (2002). Developing an assessment plan to learn about student learning. *The Journal of Academic Librarianship*, *28*(1-2), 8–13.

Malone, M. E. (2008). Training in language assessment. In N. H. Hornberger (Ed.), *Encyclopedia of language and education* (pp. 2362–2376). Springer US. https://doi.org/10.1007/978-0-387-30424-3_178

Martinez, J. G., & Martinez, N. C. (1992). Re-examining repeated testing and teacher effects in a remedial mathematics course. *British Journal of Educational Psychology*, *62*(3), 356–363.

Mattsson, M., Eilertsen, T. V., & Rorrison, D. (2011). What is practice in teacher education? In M. Mattsson, T. V. Eilertsen, & D. Rorrison (Eds.), *A practicum turn in teacher education* (pp. 1–15). Brill. https://doi.org/10.1007/978-94-6091-711-0_1

McGee, J., & Colby, S. (2014). Impact of an assessment course on teacher candidates' assessment literacy. *Action in Teacher Education*, *36*(5-6), 522–532.

McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, *20*(1), 20–32.

McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice*, *22*(4), 34–43.

McMillan, J. H., & Nash, S. (2000). *Teacher classroom assessment and grading practices decision making. (*ED447195*)*. ERIC. https://files.eric.ed.gov/fulltext/ED447195.pdf

Mehrens, W. A., & Ebel, R. L. (1979). Some comments on criterion-referenced and norm-referenced achievement tests. *NCME Measurement in Education*, *10*(1), 1–9.

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.

Mertler, C. A. (1999). Assessing student performance: A descriptive study of the classroom assessment practices of Ohio teachers. *Education*, *120*(2), 285–285.

Mertler, C. A. (2003). *Classroom assessment: A practical guide for educators*. Routledge.

Mertler, C. A. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education*, *33*(1), 49–64.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools*, *12*(2), 101–113.

Mertler, C. A., & Campbell, C. S. (2005). Measuring teachers' knowledge and application of classroom assessment concepts: Development of the assessment literacy inventory. *Annual Meeting of the American Educational Research Association*.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). Sage.

Mohono-Mahlatsi, L., & Van Tonder, F. (2006). The effectiveness of mentoring in the distance teacher education programme at the lesotho college of education: Student teachers' and tutors' perceptions. *South African Journal of Education*, *26*(3), 383–396.

Morgan, D. L. (1996). *Focus groups as qualitative research* (Vol. 16). Sage publications.

National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Author.

National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Author.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. National Academy Press.

Newfields, T. (2006). Teacher development and assessment literacy. *Authentic Communication: Proceedings of the 5th Annual JALT PanSIG Conference., 48–73.*

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in education*, *14*(2), 149–170.

Nitko, A. J. (2001). *Educational assessment of students (3rd ed.)*. Prentice-Hall, Inc.

Nitko, A. J., & Brookhart, S. M. (2007). *Educational assessment of students* (5th ed.). Pearson Merrill Prentice Hall.

Nolen, S. B., Horn, I. S., Ward, C. J., & Childers, S. A. (2011). Novice teacher learning and motivation across contexts: Assessment tools as boundary objects. *Cognition and Instruction*, *29*(1), 88–122.

Norman, A. D. (2010). Assessing accomplished teaching: Good strides, great challenges. *Theory into Practice*, *49*(3), 203–212.

Nuttall, D. L. (1987). The validity of assessments. *European Journal of Psychology of Education*, *2*(2), 109–118.

Ogan-Bekiroglu, F. (2009). Assessing assessment: Examination of pre-service physics teachers' attitudes towards assessment and factors affecting their attitudes. *International Journal of Science Education*, *31*(1), 1–39.

Ogan-Bekiroglu, F., & Suzuk, E. (2014). Pre-service teachers' assessment literacy and its implementation into practice. *The Curriculum Journal*, *25*(3), 344–371.

Organisation for Economic Co-operation and Development. (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. OECD Publishing.

O'Sullivan, R. G., & Johnson, R. L. (1993). *Using performance assessments to measure teachers' competence in classroom assessment. (*ED358156*)*. ERIC. https://https://files.eric.ed.gov/fulltext/ED358156.pdf

Otero, V. K. (2006). Moving beyond the "get it or don't" conception of formative assessment. *Journal of Teacher Education*, *57*(3), 247–255.

Özçelik, D. A. (1998). *Ölçme ve değerlendirme*. OSYM Publishing.

Parsian, A., & Rejali, A. (2011). An experience on training mathematics teachers for teaching statistics in Iran. In *Teaching statistics in school mathematics—challenges for teaching and teacher education* (pp. 37–40). Springer.

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, *84*, 128–138.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Sage Publications.

Pilcher, J. K. (2001). *The standards and integrating instructional and assessment practices. (*ED451190*)*. ERIC. https://files.eric.ed.gov/fulltext/ED451190.pdf

Plake, B. S., & Impara, J. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher*, *6*(1), 21–27.

Plake, B. S., & Impara, J. C. (1992). *Teacher competencies questionnaire description*. University of Nebraska.

Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, *12*(4), 10–12.

Popham, W. J. (2014). *Classroom assessment: What teachers need to know* (7th ed.). Pearson.

Popham, W. J. (2004). All about accountability/why assessment illiteracy is professional suicide. *Educational Leadership*, *62*(1), 82–83.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, *48*(1), 4–11.

Popham, W. J. (2018). *Assessment literacy for educators in a hurry*. ASCD.

Price, M. (2005). Assessment standards: The role of communities of practice and the scholarship of assessment. *Assessment & Evaluation in Higher Education*, *30*(3), 215–230.

Prizovskaya, V. (2017). *Identifying new jersey teachers' assessment literacy as precondition for implementing student growth objectives* (Doctoral dissertation). Grand Canyon University.

Ramsden, P. (2003). *Learning to teach in higher education*. Routledge.

Rasmussen, J., & Rash-Christensen, A. (2015). How to improve the relationship between theory and practice in teacher education. *Educational Research for Policy and Practice*, *14*(3), 213–230.

Rey, B., Carette, V., Defrance, A., Kahn, S., & Meirieu, P. (2012). *Les compétences à l'école: Apprentissage et évaluation*. De Boeck,

Reynolds, C., Livingston, R., & Willson, V. (2008). *Measurement and assessment in education* (2nd ed.). Pearson.

Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge university press.

Rockman, I. F. (2002). The importance of assessment. *Reference Services Review*, *30*(3), 181–182. https://doi.org/10.1108/00907320210435455

Russell, M. K., & Airasian, P. W. (2012). *Classroom assessment: Concepts and applications* (7th ed.). McGraw-Hill.

Rust, C. (2002a). The impact of assessment on student learning. *Active Learning in Higher Education*, *3*(2), 145–158. https : / / doi . org / 10 . 1177 / 1469787402003002004

Rust, C. (2002b). *Purpose and principles of assessment*. Oxford Centre for Staff and Learning Development.

Saad, S., Carter, G. W., Rothenberg, M., & Israelson, E. (1999). *Testing and assessment: An employer's guide to good practices*. U.S. Department of Labor.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77–84.

Sakash, K., & Rodriguez-Brown, F. (2010). Fostering collaboration between mainstream and bilingual teachers and teacher candidates. In T. Lucas (Ed.), *Teacher preparation for linguistically diverse classrooms: A resource for teacher educators* (pp. 143–159). Routledge.

Sanders, W. L., & Horn, S. P. (1995). Educational assessment reassessed. *Education policy analysis archives*, *3*(6), 1–15. https://doi.org/10.14507/epaa.v3n6.1995

Schafer, W. D. (1993). Assessment literacy for teachers. *Theory into Practice*, *32*(2), 118–126.

Schwab, S., Resch, K., & Alnahdi, G. (2021). Inclusion does not solely apply to students with disabilities: Pre-service teachers' attitudes towards inclusive schooling of all students. *International Journal of Inclusive Education*, *0*(0), 1–17. https://doi.org/10.1080/13603116.2021.1938712

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Rand McNally.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63–75.

Shepard, L. (1979). Norm-referenced vs. criterion-referenced tests. *Educational Horizons*, *58*(1), 26–32.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4–14.

Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education*, *22*(4), 371–391.

Simeonsson, R. J., & Bailey, D. B. (1988). Essential elements of the assessment process. In T. D. Wachs & B. Sheehan (Eds.), *Assessment of developmentally disabled infants and preschool children* (pp. 25–41). Springer.

Stake, R. E. (2006). *Multiple case study analysis*. The Guilford Press.

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, *72*(7), 534–539.

Stiggins, R. J. (1994). *Student-centered classroom assessment*. Merrill New York.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, *77*(3), 238.

Stiggins, R. J. (1999). Assessment, student confidence, and school success. *The Phi Delta Kappan*, *81*(3), 191–198.

Stiggins, R. J., & Conklin, N. F. (1988). *Teacher training in assessment (*ED303439*)*. ERIC. https://files.eric.ed.gov/fulltext/ED303439.pdf

Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, *86*(1), 22–27.

Stiggins, R. (2005). From formative assessment to assessment for learning: A path to success in standards-based schools. *Phi Delta Kappan*, *87*(4), 324–328.

Stiggins, R., & Chappuis, J. (2008). Enhancing student learning. *District Administration*, *44*(1), 42–44.

Strauss, A., & Corbin, J. (2004). Open coding. In C. Seale (Ed.), *Social research methods: A reader* (pp. 303–306). Sage Publications.

Sugianto, A. (2020). Applying a lesson plan for a digital classroom: Challenges and benefits. *International Journal of English Education and Linguistics*, *2*(2), 21–33.

Suurtamm, C., Koch, M., & Arden, A. (2010). Teachers' assessment practices in mathematics: Classrooms in the context of reform. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 399–417.

Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., Silver, E., Ufer, S., & Vos, P. (2016). *Assessment in mathematics education: Large-scale assessment and classroom assessment*. Springer Nature.

Taylor, C. S., & Nolen, S. B. (2008). *Classroom assessment: Supporting teaching and learning in real classrooms* (2nd ed.). Pearson/Merrill/Prentice Hall.

Thomas, J., Allman, C., & Beech, M. (2004). *Assessment for the diverse classroom: A handbook for teachers*. Florida Department of Education, Bureau of Exceptional Education; Student Services.

Tsagari, D. (2016). Assessment orientations of state primary EFL teachers in two Mediterranean countries. *Center for Educational Policy Studies Journal*, *6*(1), 9–30.

Tünkler, V. (2019). Sosyal bilgiler öğretmen adaylarının değerlendirme okuryazarlık düzeyi ile değerlendirme öz-yeterlik algılarının İncelenmesi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (37), 319–332.

Turgut, M. F., & Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme*. Pegem Akademi.

Uludag, P., & McDonough, K. (2022). Validating a rubric for assessing integrated writing in an EAP context. *Assessing Writing*, *52*, 1–12.

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education*, *30*(3), 749–770.

Wagner, D. (2009). New technologies for adult literacy and international development. In D. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 558–565). Cambridge University Press.

Wiliam, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F. K. Lester (Ed.), *Second handbook of mathematics teaching and learning* (pp. 1053–1098). Information Age Publishing.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3–14.

Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education*, *42*(1), 37–42.

Xu, Y. (2019). English language teacher assessment literacy in practice. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 517–539). Springer.

Xu, Y., & Brown, G. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, *58*, 149–162.

Yamtim, V., & Wongwanich, S. (2014). A study of classroom assessment literacy of primary school teachers. *Procedia-Social and Behavioral Sciences*, *116*, 2998–3004.

Yan, X., Zhang, C., & Fan, J. J. (2018). "Assessment knowledge is important, but. . . ": How contextual and experiential factors mediate assessment practice and training needs of language teachers. *System*, *74*, 158–168.

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users*. Sage Publications.

Ye, W. (2023). Urban public-school teachers' assessment literacy in China's Zhengzhou city. *Education 3-13*, *51*(5), 836–849.

Yilmaz-Tuzun, O. (2008). Preservice elementary teachers' beliefs about science teaching. *Journal of Science Teacher Education*, *19*(2), 183–204.

Yıldırım, A., & Şimşek, H. (2013). *Sosyal bilimlerde nitel araştırma yöntemleri* (9th ed.). Seçkin.

Yılmaz, H. (2002). *Eğitimde ölçme ve değerlendirme* (5th ed.). Çizgi Kitabevi.

Young, G. (1999). Using portfolios for assessment in teacher preparation and health sciences. In S. Brow & A. Glasner (Eds.), *Assessment matters in higher education* (pp. 122–131). Open University Press.

Yuan, R., & Zhang, J. (2016). Promoting teacher collaboration through joint lesson planning: Challenges and coping strategies. *The Asia-Pacific Education Researcher*, *25*, 817–826.

Yüksek Öğretim Kurulu (YÖK). (2007). *Eğitim fakültesi öğretmen yetiştirme programları*. Yükseköğretim Kurulu Yayını.

Zembal-Saul, C., Starr, M. L., & Krajcik, J. S. (1999). Constructing a framework for elementary science teaching using pedagogical content knowledge. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 237–256). Springer.

Zolfaghari, F., & Ahmadi, A. (2016). Assessment literacy components across subject matters. Cogent Education, 3(1), 1–16.

Appendix A

MSE 305: ASSESSMENT OF LEARNING IN SCIENCE AND MATHEMATICS COURSE SYLLABUS (FALL 2022-2023)

Catalog Description

Focuses on construction and use of classroom tests to assess student learning in relation to instructional objectives, test interpretation, basic psychometric statistics and reporting.

Course Objectives and Goals

You are expected to achieve the following goals:

- understand basic concepts related to assessment and measurement.
- understand the role of measurement and assessment in the instructional process.
- identify instructional objectives as intended learning outcomes.
- understand the issues that make a well-designed instrument for classroom evaluation.
- develop various assessment materials for classroom use.

Specifically, we will focus on establishing a framework for assessing students, types of instructional decisions, validity, and reliability issues, developing, administering, and scoring written tests and alternative assessments, and using assessment for instructional decisions.

Supplemental Textbooks

- Taylor, C. S. & Nolen, S. B. (2008). Classroom assessment: Supporting teaching and learning in real classrooms (2nd ed.). Upper Saddle River, NJ: Pearson.

- Thorndike-Christ, T. (2014). Measurement and evaluation in psychology and education (New international ed.). Edinburg Gate: Pearson.

Tentative Schedule

Tentative Schedule

Week	Date	Topic	Readings/Assignments Due
1	Oct 3	Introduction	
2	Oct 10	How to use Assessment	• Week2_ks4_guidance_on_teaching • Formative Assessment: ○ http://map.mathshell.org/background.php?subpage=formative ○ http://map.mathshell.org/pd/modules/1_Formative_Assessment/html/index.htm • Summative Assessment: ○ http://map.mathshell.org/background.php?subpage=summative ○ http://map.mathshell.org/tasks.php?gradeid=25
3	Oct 17	Validity	
4	Oct 24	Generalizing Observed Performance	
5	Oct 31	Completion and Short-Answer	
6	Nov 8	Essay Items	
7	Nov 15	Multiple-Choice	
8	**Nov 22**	**MIDTERM I**	
9	Nov 29	Alternate-Choice Items	
10	Dec 6	Performance Assessments	
11	Dec 13	Informal Assessment Portfolios	
12	Dec 20	Norm-Referenced & Standards-Based Test Scores	
13	**Dec 27**	**MIDTERM II**	
14	Jan 3	Statistics of the Measurement	• statistics worksheet
FINAL Exam (The date, time, and details for the final exam will be announced.)			

Appendix B

ASSESSMENT LITERACY INVENTORY

Değerli Öğretmenler /Öğretmen Adayları,

Değerli cevaplarınıza ihtiyaç duyulan bu envanter, öğretmen adaylarının değerlendirme alanındaki bilgi ve becerilerini yansıtan "Değerlendirme Okuryazarlığı" düzeylerini anlamaya yönelik uygulanmaktadır. Bu envanter, her biri altı sorudan oluşan toplam beş senaryodan oluşmaktadır. Toplamda 30 sorudan bu envanteri tamamlamak yaklaşık 40 dakika sürmektedir. Senaryoları ve bu senaryolara ait soruları dikkatlice okumanız, ardından en uygun olduğunu düşündüğünüz yanıtı işaretlemeniz araştırmanın bilimsel olarak geçerli olması açısından büyük önem taşımaktadır. Seçiminizden emin olmasanız bile, soruları boş bırakmamanız ve en uygun cevap olduğuna inandığınız şıkkı işaretlemeniz önemlidir.

Araştırmanın devamında öğretmen adaylarıyla görüşmeler yapılacağı için demografik bilgiler bölümünde katılımcıların isim ve soy isimlerini yazacakları bir alan bırakılmıştır. Araştırmanın amacına ulaşması için büyük katkı sağlayacak olan cevaplarınız ve bilgileriniz, sadece araştırma için kullanılacak ve başka bir amaç için kimseyle paylaşılmayacaktır.

Çalışmaya verdiğiniz katkılarınız için teşekkür ederim.

Beyza ŞAHİN

ODTÜ Matematik ve Fen Bilimleri Eğitimi Bölümü Yüksek Lisans Öğrencisi

<u>BÖLÜM 1:</u> Demografik Bilgiler (Pre-service Teachers)

Lütfen, durumuza en uygun olan seçenekleri yanındaki boşluğa çarpı işareti (X) koyarak belirtiniz.

1. İsminiz ve soy isminiz: .....

2. Yaşınız: .....

3. Cinsiyetiniz: ( ) Kadın ( ) Erkek

4. Sınıf Seviyeniz: ( ) 3. Sınıf ( ) 4. Sınıf

5. Akademik başarı ortalamanız: ( ) 3.51 – 4.00
( ) 3.01 – 3.50
( ) 2.51 – 3.00
( ) 2.00 – 2.50
( ) 2.00'den düşük

4. Aşağıdaki derslerden aldıklarınızı veya alıyor olduklarınızı lütfen işaretleyiniz.

( ) STAT 201: Introduction to Probability &Stat. I
( ) STAT 202: Introduction to Probability &Stat. II
( ) MSE 305: Assessment of Learning in Science and Mathematics
( ) MSE 341: Methods of Teaching Mathematics In Middle Schools I
( ) MSE 342: Methods of Teaching Mathematics In Middle Schools II
( ) MSE 441: Methods of Teaching Mathematics In Middle Schools III
( ) MSE417: School Experience
( ) MSE418: Practice Teaching

5. Değerlendirme ile ilgili zorunlu dersler dışında herhangi eğitim/faaliyet/etkinlik aldınız mı?

( ) Evet ( ) Hayır

*Eğer cevabınız **evet** ise lütfen 6. soruyu da cevaplayınız.*

6. Değerlendirme ile ilgili aldığınız eğitim/faaliyet/etkinlik isimlerini yazınız: .....

7. Çalışmanın devamında yapılacak olan görüşmeye katılmak ister misiniz?

( ) Evet ( ) Hayır

<u>BÖLÜM 1:</u> Demografik Bilgiler (In-service Teachers)

Lütfen, durumuza en uygun olan seçenekleri yanındaki boşluğa çarpı işareti (X) koyarak belirtiniz.

1. İsminiz ve soy isminiz: ……..

2. Yaşınız: ……

3. Cinsiyetiniz: ( ) Kadın ( ) Erkek

4. Hangi okul ve bölümden mezunsunuz?: ……..

5. Öğrenim durumunuz: ( ) Lisans ( ) Yüksek Lisans ( ) Doktora

6. Meslekteki hizmet yılınız: ( ) 0 – 2 ( ) 3 – 5 ( ) 6 – 9 ( ) 10 – 14 ( ) 15 ve yukarı

7. Daha önce görev yaptığınız okul türlerini işaretleyiniz: ( ) Özel ( ) Devlet

8. Şu an görev yaptığınız okulun türü: ( ) Özel ( ) Devlet

9. Şu an görev yaptığınız okulun bölgesi: ( ) Kırsal ( ) Şehir merkezi

10. Üniversitede ölçme-değerlendirme dersi aldınız mı? : ( ) Evet ( ) Hayır

11. Değerlendirme ile ilgili zorunlu dersler dışında herhangi eğitim/faaliyet/etkinlik aldınız mı?: ( ) Evet ( ) Hayır

Eğer cevabınız evet ise lütfen 12. soruyu da cevaplayınız.

12. Değerlendirme ile ilgili aldığınız eğitim/faaliyet/etkinlik isimlerini yazınız: ……..

13. Çalışmanın devamında yapılacak olan görüşmeye katılmak ister misiniz?: ( ) Evet ( ) Hayır

Senaryo 1

Matematik öğretmeni Ayşe Hanım, 10. sınıftaki öğrencilerinin sınıfta öğrendiklerini gerçek hayatta ne kadar uygulayabildiklerini görmek istemektedir. Öğretmen kılavuz kitabı öğrencilerin matematiksel kavramları anlama düzeylerini ölçmek için çok sayıda yazılı sınav sorusu içermesine rağmen, Ayşe Hanım istediğini yapmak için yazılı sınavların en iyi yöntem olduğuna inanmamaktadır.

1. Yukarıdaki senaryoya göre, Ayşe Hanım'ın amacına **en uygun ölçme-değerlendirme** şekli aşağıdakilerden hangisidir?

A) Performans değerlendirme
B) Özgün (authentic, gerçek) değerlendirme
C) Uzun cevaplı sorular
D) Genel testler

2. Öğrencilerinin bilgilerini doğru ve tutarlı bir şekilde değerlendirmek için Ayşe Hanım'a aşağıdakilerden hangisi önerilebilir?

A) Ünite hedeflerine uygun ölçütler belirlemeli ve bir puanlama ölçeği (rubric) geliştirmeli.
B) Öğrencilerin ne yapabileceğini belirledikten sonra bir puanlama ölçeği geliştirmeli.
C) Öğrencilerin benzer görevlerdeki performanslarını göz önüne almalı.
D) Geçmişte kullanılmış ölçütler hakkında deneyimli meslektaşlarının görüşünü almalı.

3. Ayşe Hanım öğrencilerinin performansının diğer 10. Sınıf öğrencilerine göre ne durumda olduğu hakkında genel bir izlenim edinmek için, genel bir matematik testi uyguluyor. Bu uygulama, aşağıdaki şartlardan hangisi yerine geldiğinde kabul edilebilir?

A) Testin güvenirliği 0,60'ı geçmezse
B) Genel matematik testi öğrenciler tarafından bireysel olarak cevaplanırsa
C) Testte sorulan konular öğrenciler tarafından iyi bilinirse
D) Karşılaştırma yapılacak kontrol grubu aynı sınıf düzeyinde öğrencilerden oluşursa

4. Aşağıdaki durumlardan hangisi, 3. sorudaki matematik testinden alınan sonuçların kullanımı için **uygun değildir**?

A) Öğretimi planlamak
B) Öğrencilere not vermek
C) Öğrencilerin güçlü ve zayıf yönlerini belirlemek
D) Öğretim programı (curriculum) geliştirmek

5. Ayşe Hanım öğrencilerinin dersi anlama düzeylerini öğretim boyunca değerlendirmektedir. Bu değerlendirmede, yeni konunun öğretimini izleyen kısa sınavlardan, ünite sonu sınavlarına kadar değişen ölçme-değerlendirme yöntemlerini kullanmaktadır. Ayşe Hanım yaptığı not verme işleminin geçerliğini artırmak için ne yapmalıdır?

A) Bütün değerlendirmeler için aynı derecelendirme ölçeğini kullanmalıdır.
B) Final notunu verirken öğrencilerin önceki performanslarını dikkate almalıdır.
C) Yaptığı farklı değerlendirmelerin önemine göre öğrenci notlarındaki ağırlığı belirlemelidir.
D) Notları hesaplarken, öğrencilerin gösterdiği çabayı dikkate almalıdır.

6. Veli toplantısı sırasında Ayşe Hanım'ın sınıfındaki öğrencilerden birinin velisi, kızının matematik puanının 80. yüzdelikte olmasının ne anlama geldiğini öğrenmek istemiştir. Bu öğrencinin puanı için **en iyi açıklama** aşağıdakilerden hangisidir?

A) Testteki soruların %80'ini doğru cevaplamıştır.
B) Sınıfında 5 üzerinden 4 alacak düzeyde başarı göstermiştir.
C) Matematikte sınıf seviyesinin üzerinde bir başarı göstermiştir.
D) Sınıfın %80'inden daha iyi puan almıştır.

Senaryo 2

Kemal Bey, 5. sınıf öğretmenidir. Gelecek yıl yapacağı öğretimi planlamaktadır. Aynı zamanda, öğrencilerinin gelecek yılın sonunda il genelinde yapılacak bir başarı sınavına gireceğinin de farkındadır.

7. Kemal Bey'in bu yılki matematik ünitesi çok adımlı problem çözmeye odaklanacaktır. Kemal Bey, il genelinde yapılacak sınavdan önce bazı konuların tekrar anlatılmasının gerekli olup olmadığını anlamak için ünitenin sonunda öğrencilerinin problem çözme becerilerini ölçmeye yönelik bir sınav yapmak istemektedir. Aşağıdaki ölçme-değerlendirme yöntemlerinden hangisi bu amaç için **en uygun seçim** olacaktır?

A) Ders kitabının öğretmen kılavuzunda yer alan bir yöntem
B) Öğretmiş olduğu beceri ve içerikle uyumlu bir yöntem
C) Standart ölçme değerlendirme yöntemi
D) Tek adımlı problem çözme yeteneklerini içeren bir yöntem

8. Kemal Bey, öğrencileri için tekrar öğretimin gerekli olup olmadığını belirlemek için, kendi ölçme aracını geliştirmeye karar verir. Kemal Bey, bu testi aynı zamanda öğrencilerinin genel değerlendirme testinde nasıl bir performans göstereceğini de tahmin etmek için kullanmak istemektedir. Öğrencilerinin performansını en doğru şekilde tahmin etmek için, Kemal Bey'in geliştireceği **en uygun değerlendirme şekli** aşağıdakilerden hangisidir?

A) Performans değerlendirmesi
B) Çoktan seçmeli test
C) Portfolyo değerlendirmesi
D) Açık uçlu sorular

9. Kemal Bey'in öğrencisi Elif, il genelinde yapılan testin problem-çözme becerileri bölümünde %60'lık dilime girmiştir. Bu sonuç **en iyi** aşağıdakilerden hangisi ile yorumlanabilir?

A) Elif ortalamanın üzerinde puan almıştır.
B) Elif ortalamanın altında puan almıştır.
C) Elif ülke ortalamasında puan almıştır.
D) Yorum yapmak için yeterli bilgi yoktur.

10. Ahmet, Kemal Bey'in başka bir öğrencisidir. Ahmet, testin okuma bölümünden 196 puan almıştır. Geçebilmek için sınır değer 200'dür. Sonuçta Ahmet testten geçememiştir. Bununla birlikte, testin standart hatası 6 olarak ölçülmüştür. Ahmet'in ihtiyaçlarına cevap verecek öğretimi gerçekleştirmede aşağıdakilerden hangisi Kemal Bey için **en iyi karar** olacaktır?

A) Ahmet'in okuma becerisi için gerekli minimum puanı alamadığı açıktır bu yüzden okuma becerisini geliştirmeye yönelik ek ders almalıdır.
B) Kemal Bey, Ahmet'in daha yüksek alabileceğini bilmektedir, bu yüzden testin sonuçları göz ardı edilebilir.
C) Ahmet minimum puanı az farkla kaçırdığı için fazladan veya farklı bir şey yapmaya gerek yoktur.
D) Kemal Bey, Ahmet'in aslında çok daha düşük puan alması gerektiğini bildiği için test sonuçları göz ardı edilmelidir.

11. Kemal Bey'in uygulamayı düşündüğü aşağıdaki yöntemlerden hangisi ile verilen not başarıyı **en az yansıtır**?

A) Günlük ödevler ve bölüm testleri
B) Düşük performansta notun kırıldığı günlük ödevler ve ünite testleri üzerine verilen notlar
C) Öğrencilerin daha yüksek başarı sağlayabilmeleri amacıyla yeniden yapmalarına izin verilen günlük ödevler ve ünite sonu testleri
D) Günlük ödevlerin resmi olarak değerlendirmeye alınmadığı ünite testlerine dayalı verilen notlar

12. Neslihan'ın matematik problem çözme testinden aldığı puan 60. yüzdelik sırasında ve okuduğunu anlama testinden aldığı puan ise 56. yüzdelik sırasında yer almaktadır. Her iki test için de yüzdelik dilimlerin genişliği %5'tir. Kemal Bey, Neslihan'ın ailesine nasıl bir tavsiyede bulunmalıdır?

A) Farklılığı görmezden gelmeliler; öğrencinin performansı iki testte de aynıdır.
B) Neslihan'ın okumasını geliştirmesi için ek ders aldırmalılar.
C) Neslihan'ı evde daha çok okumaya zorlamalılar.
D) Neslihan için daha başarılı olduğu matematikte gelişmesi için destek olmalılar.

Senaryo 3

Fatma Hanım, 8. sınıf tarih öğretmenidir. Ünitenin sonunda, öğrencilerinin üst düzey düşünme becerilerini belirlemek için çoktan seçmeli bir test uygulamak istemektedir. Öğrencilerinin çoğunun bu sınavda başarılı olacağını düşünmektedir.

13. Fatma Öğretmen'in amacını dikkate alarak, çoktan seçmeli test uygulama yönünde verdiği bu kararla ilgili ne söylenebilir?
- A) Ünitenin değerlendirilmesi için uygun bir seçimdir.
- B) Test puanları amaç için geçerli (valid) olmayabilir.
- C) Test puanları amaç için güvenilir (reliable) olmayabilir.
- D) Doğru/yanlış testi daha uygun olacaktır.

14. Yaptığı çoktan seçmeli testin kalitesini belirlemek için, Fatma Öğretmenin uygulayacağı madde analizi aşağıdakilerden **hangisini içermeyecektir**?
- A) Madde güçlük indeksi (item difficulty values)
- B) Madde ayırt edicilik indeksi (item discrimination values)
- C) Güvenirlik katsayısı (reliability coefficient)
- D) Geçerlilik katsayısı (validity coefficients)

15. Fatma Öğretmen, sınav kâğıtlarını 100 üzerinden puanlamaya karar verir. Buna göre genel olarak, bir öğrencinin 85 alması nasıl yorumlanabilir?
- A) Öğrenci test sorularının %85'ini doğru yanıtlamıştır
- B) Öğrenci ünite içindeki konuların %85'ini bilmektedir.
- C) Öğrenci sınava giren öğrencilerin %85'inden daha yüksek puan almıştır.
- D) Öğrenci sınava giren öğrencilerin %85'inden daha düşük puan almıştır.

16. Fatma Öğretmenin öğrencilerinden bazıları çoktan seçmeli testten iyi puanlar alamamışlardır. Fatma Öğretmen ilerde aynı üniteyi tekrar işleyeceği zaman öğrencilerinin hazır bulunuşluk düzeylerini belirlemek için bir ön test kullanmaya karar verir. Bu testin sonucuna göre öğretimini planlayacaktır. Fatma Öğretmen burada ne tür bir değerlendirme yapmaktadır?
- A) Bağıl Değerlendirme (norm-referenced information)
- B) Mutlak Değerlendirme (criterion-referenced information)
- C) Hem mutlak hem bağıl değerlendirme
- D) Hiç birisi

17. Fatma Öğretmen, öğrencilerini bu dönemde sadece bir konudan test yapmıştır. Bu yüzden öğrencilerin notları sadece bu test puanlarıyla belirlenmiştir. Bu uygulamanın **temel eleştirisi** ne olabilir?
- A) Bu test öğrencileri tüm programa göre değerlendirmede çok sınırlı kalmıştır.
- B) Sadece teste dayalı değerlendirme olduğundan, bazı öğrenciler için dezavantajlı olabilir.
- C) Testten düşük puan alan öğrencilerin puanlarına ekleme yapmalıdır.
- D) Öğrencilerin notları konusundaki kararlar birden çok bilgiye dayanmalıdır.

18. İsmail Bey, diğer bir tarih öğretmenidir. İsmail Bey, notlarını öncelikli olarak sınıftaki gözlemlerine dayalı olarak vermektedir. İsmail Bey ile Fatma Hanım'ın öğrencilerini değerlendirmedeki temel farklılığı aşağıdakilerden hangisi **en iyi şekilde açıklamaktadır**?
- A) Fatma Öğretmen formal değerlendirmeyi, İsmail Öğretmen informal değerlendirmeyi kullanmaktadır.
- B) Fatma Öğretmen biçimlendirici (formative) değerlendirmeyi, İsmail Öğretmen geleneksel (summative) değerlendirmeyi kullanmaktadır.
- C) Fatma Öğretmen standart değerlendirmeyi, İsmail Öğretmen standart olmayan değerlendirmeyi kullanmaktadır.
- D) Fatma Öğretmen geleneksel değerlendirmeyi, İsmail Öğretmen alternatif değerlendirmeyi kullanmaktadır.

Senaryo 4

Ali Bey, yeni açılan bir ilköğretim okulunda Türkçe öğretmenidir. Sınıf değerlendirmesi konusunda tecrübeli olan Ali Bey'e sık sık etkili değerlendirme yöntemleri konusunda danışılmaktadır.

19. Leyla Hanım, okulun diğer bir Türkçe öğretmenidir. Ali Bey'e, 6. sınıfların yazma yeteneklerini değerlendirmede en iyi yolun ne olduğunu sorar. Bu soruya **en iyi cevap** aşağıdakilerden hangisidir?

A) Çoktan seçmeli sorular
B) Doğru / Yanlış ifadeleri
C) Boşluk doldurma soruları
D) Kompozisyon yazma

20. İlköğretim matematik öğretmenlerinden biri, öğrencilerinin matematiği anlama düzeylerini ölçmede bir yol olarak hikâye tarzı problemlere daha çok yer vermek için testlerini yeniden düzenlemektedir. Bu tür testleri hazırlarken karşılaşabileceği sıkıntılar konusunda Ali Bey'in fikrini almak ister. Aşağıdakilerden hangisi hikâyeye dayalı matematik testlerinin hazırlanması konusunda uygun bir öneri **olamaz**?

A) Okuma düzeyi sınıf seviyesine uygun olmalıdır.
B) Bazı grup öğrencilerine diğerlerinden daha tanıdık gelebilecek senaryolardan kaçınılmalıdır.
C) Cümlelerin anlaşılır olup olmadığı kontrol edilmelidir.
D) Derste kullanılmış senaryolardan yararlanılmalıdır.

21. Ali Öğretmen'in sınıfının bir öğrencisi olan Esra, standart sapmanın 4 ve ortalamanın 80 olduğu bir Türkçe başarı testinden 78 puan almıştır. Bu testin standart sapması 3 ve ortalaması 50 olan fen bölümünden ise 60 puan almıştır. Yukarıdaki bilgilere göre, akranlarıyla karşılaştırıldığında, Esra'nın durumu ile ilgili olarak **en doğru** ifade aşağıdakilerden hangisinde verilmiştir?

A) Esra Türkçe 'de, Fen'e göre daha iyidir.
B) Esra Fen'de, Türkçeye göre daha iyidir.
C) Esra her iki derste de sınıf ortalamasının altında kalmıştır.
D) Esra her iki derste de sınıf ortalamasına yakın seviyededir.

22. Ali Öğretmen her dersin sonunda, öğrencilerinin anlama düzeylerini kontrol etmektedir. Bu şekilde, biçimlendirici (formative) ölçme-değerlendirme yapmanın **temel amacı** aşağıdakilerden hangisidir?

A) Bilgideki artışı takip etmek
B) Final sınavının içeriğini belirlemek
C) Dersi planlamak
D) Öğretim programının (curriculum) uygunluğunu değerlendirmek

23. Öğrencileri genel sınava hazırlamak ve okul gelişim alanlarını belirlemek için tüm 6. sınıf Türkçe öğretmenleri bir dizi açık uçlu soru içeren ortak bir final sınavı yaparlar. Ancak son zamanlarda bazı öğretmenler, sınav kâğıtlarını okumayı zamanında yetiştirme kaygısıyla bazı tutarsız sonuçların doğabileceği yönünde endişeler bildirmişlerdir. Bu konuda Ali Öğretmen'e danışırlar. Aşağıdakilerden hangisi öğretmenlerin tutarlılık konusundaki endişesine verilebilecek **en iyi cevaptır**?

A) Tüm öğrenciler için önce birinci soruları, sonra ikinci soruları olmak üzere sırayla okuyun.
B) Puanlama esnasında puanlama anahtarında örnek öğrenci çalışmalarını gösterecek şekilde düzenlemeler yapın.
C) Puanlamanın öğretmene göre değişmesini minimum düzeye indirmek için bütüncül (holistic) puanlama yöntemi kullanın.
D) Klasik yazılı sınavları daha az kullanın.

24. Utku, Ali Öğretmen'in sınıfındaki 6. sınıf öğrencisidir. Genel okuma testinde 0,5'e eşdeğer bir z puanı almıştır. Utku'nun ailesi bunun ne anlama geldiğini merak etmektedir. Yukarıdaki bilgiye göre, öğrencinin puanı konusunda **en uygun yorum** aşağıdakilerden hangisidir?

A) Utku 7. sınıf düzeyinde okuyor.
B) Utku sınıfındaki öğrencilerin çoğundan daha iyi okuyor.
C) Utku beklendiği gibi 6. sınıf düzeyinde okuyor.
D) Utku 7. sınıf okuma sınıfına yerleştirilmelidir.

Senaryo 5

Suzan Hanım 6. sınıf Fen Bilgisi öğretmenidir. Son yıllarda öğrencilerinin hal değişimi (donma, erime, yoğunlaşma, buharlaşma vb.) konusunu anlamakta bazı zorluklar yaşadığını fark etmekle birlikte, özellikle nerelerde sorun yaşadıkları konusunda emin değildir. Ancak öğrencilerinin bu konuda gelişmeleri gerektiğini düşünmektedir.

25. Suzan Öğretmen, öğrencilerinin özellikle hangi noktalarda zorluk yaşadıklarını belirlemek istemektedir. Aşağıdakilerden hangisi onun ihtiyaçlarına **en iyi şekilde karşılık** verir?

A) Teşhis edici ölçme-değerlendirme (a diagnostic assessment)
B) İnformal ölçme-değerlendirme (an informal assessment)
C) Standart ölçme- değerlendirme (a standardized assessment)
D) Geleneksel ölçme-değerlendirme (a summative assessment)

26. Suzan Öğretmen hal değişimi konusunun öğretimi ve değerlendirmesini daha sağlıklı bir şekilde yapmak için, geçen yılın son sınavında uygulanan testte bu konu ile ilgili madde analizi (item analysis) yapmıştır. Aşağıdakilerden durumların hangisinde öğretmen test maddesini tekrar gözden geçirmeli veya tamamen çıkarmalıdır?

A) Zorluk indeksi (difficulty value) 0,50 ile 0,75 arasında olduğunda
B) Ayırt edicilik indeksi (discrimination value) + 0,30 olduğunda
C) Ayırt edicilik indeksi - 0,50 olduğunda
D) Zorluk indeksi 0,90 olduğunda

27. Suzan Öğretmen'in uyguladığı ünite sonu testi kısa cevaplı ve açık uçlu sorular da içermektedir. Suzan Öğretmen, öğrencilerinin cevaplarında belirli bazı kriterleri anlama düzeylerini belirlemeye çalışmaktadır. Öğrenci cevaplarının puanlanmasını aşağıdakilerden hangisi **en fazla kolaylaştırır**?

A) Objektif Cevap Anahtarı
B) Bütüncül Puanlama Ölçeği (a holistic rubric)
C) Kontrol Listesi
D) Analitik Puanlama Ölçeği (an analytic rubric)

28. Ünite sonunda, Suzan Öğretmen öğrencilerinin üniteyle ilgili kavramları kendilerinden beklenen düzeyde anladıklarını belirlemiştir. Ancak, öğrenciler ilkbaharda genel bir teste tabi tutulduklarında, öğrencilerin bu üniteyle ilgili kavramları içeren maddelerde düşük bir performans gösterdiklerini görmüştür. Öğrencilerin sınıf performansları ve genel test sonuçları arasındaki uyuşmazlık düşünülürse, okul gelişimi ile ilgili kararlar alınırken yapılması gereken **en uygun** hareket hangisidir?

A) Sınıftaki öğretimin tüm 6. sınıf öğretmenleri arasında tutarlı hale getirilmesi yönünde öneride bulunmak
B) Merkezil testlerde sorulanlarla sınıfta öğretilenlerin birbiri ile uyumunu sağlamak
C) Fen bilgisinde daha yüksek puanların alınabileceği bir test seçmek
D) İleri fen sınıflarında başarılı olması beklenen öğrencilerin yüzdesini belirlemek

29. Suzan Öğretmen, fen bilgisi dersinde öğrencilerine verdiği dönem notlarının, öğrencilerin üniteye ait içeriğe hâkimiyet düzeylerini yansıtacağından emin olmak istemektedir. Aşağıdaki not verme sistemlerinden hangisi bu amacı **en iyi** şekilde gerçekleştirir?

A) Mutlak değerlendirme (a criterion-referenced grading system)
B) Bağıl değerlendirme (a norm-referenced grading system)
C) Geçti-Kaldı şeklinde not verme
D) Portfolyo değerlendirme

30. Nalân, Suzan öğretmenin sınıfındaki bir öğrencidir. Genel bir testin fizik bölümündeki 15 sorudan 12'sini doğru cevaplamıştır. Bu ham puan %45. sıraya karşılık gelmektedir. Ailesi bu kadar soruyu doğru cevaplamasına karşın nasıl bu kadar düşük bir yüzde sırası almasına şaşırmıştır. Suzan Öğretmen'den bir açıklama beklemektedirler. Nalân'ın ailesine yapılabilecek en uygun açıklama aşağıdakilerden hangisidir?

A) Bilmiyorum. Soruların değerlendirilmesiyle ilgili bir sorun olmalı.
B) Nalân 12 tane doğru yapmış olabilir ama demek ki çok fazla sayıda çocuk 12'den fazla soruyu doğru yapmış.
C) Ham puanlar tamamen mutlak değerlendirme ile yüzdelik sıralar ise bağıl değerlendirme ile belirlenir.
D) Ham puanlar tamamen bağıl değerlendirme ile, yüzdelik sıralar ise mutlak değerlendirme ile belirlenir.

Appendix C

INTERVIEW PROTOCOL

Part 1: Beginning

1. What is the importance of measurement and assessment in mathematics education?
    (a) Which measurement and assessment tools or strategies do you use?
        i. How often do you use these tools and strategies for measurement and assessment?

Part 2: Deficiencies (Questions for pre-service teachers)

1. Do you feel competent in choosing appropriate assessment methods? Why?
    (a) How did the teacher education program fall short in your choosing appropriate assessment methods?
        i. How did the assessment/method/internship/statistic courses you took in your undergraduate education fall short in choosing appropriate assessment methods?

Part 2: Deficiencies (Questions for in-service teachers)

1. Do you feel competent in choosing appropriate assessment methods? Why?
    (a) When you became a teacher, was the information you gained from your undergraduate education sufficient to choose appropriate assessment methods?

(b) How did the teacher education program fall short in your choosing appropriate assessment methods?

i. How did the assessment/method/internship/statistic courses you took in your undergraduate education fall short in choosing appropriate assessment methods?

(The provided examples here only include the first competency, "choosing appropriate assessment methods." However, In Part 2, which addresses the deficiencies in assessment literacy of the teacher training program, the same questions have been asked separately for each competency area.)

Part 3: Suggestions

1. What are your suggestions for making the in-class and out-of-class activities and experiences more quality and effective in assessment?

   (a) What would be more effective for you if the content of the in-class experiences you took in your undergraduate education was?

      i. What could have changed or added to the content of the assessment/method/internship/statistic course? (Lesson plans, examine students' assessment data etc.)

      ii. Is the number or content of the elective courses offered in the teacher training program for measurement and assessment sufficient? What would you suggest doing about this issue?

   (b) What experiences outside the teacher education program will help you improve your measurement and assessment?

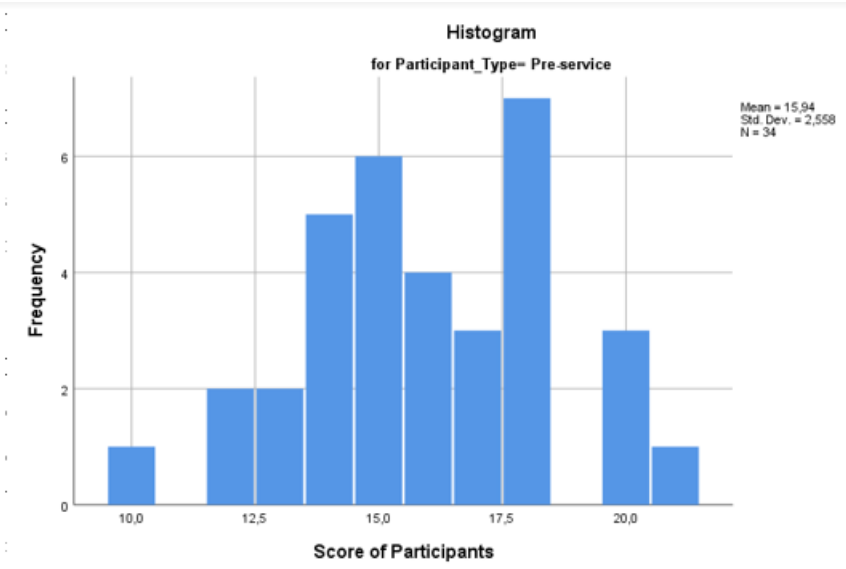
      i. What experiences can your university provide for undergraduate students outside the teacher education program?

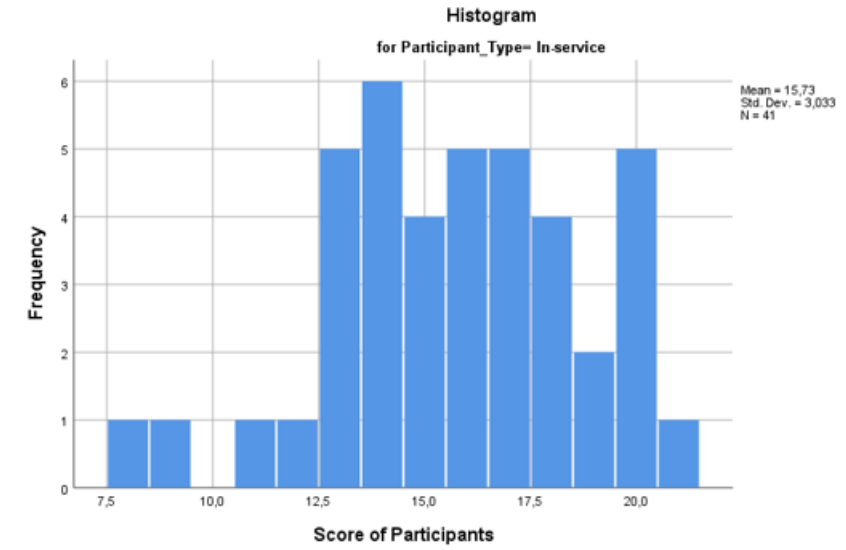
      ii. What kind of experience can other public or private institutions and organizations provide to assist your development in measurement and evaluation?

   (c) How can you develop yourself individually in measurement and assessment?

Appendix D

NORMALITY HISTOGRAMS





Appendix E

CONSENT FORM

ARAŞTIRMAYA GÖNÜLLÜ KATILIM FORMU

Bu araştırma, Orta Doğu Teknik Üniversitesi, Eğitim Fakültesi, Matematik Eğitimi programında tezli yüksek lisans öğrencisi Beyza Şahin ve danışmanı Doç. Dr. Bülent Çetinkaya tarafından yürütülmektedir. Bu form sizi araştırma koşulları hakkında bilgilendirmek için hazırlanmıştır.

Çalışmanın Amacı Nedir?

Araştırmanın iki ana amacı bulunmaktadır: Orta Doğu Teknik Üniversitesi, Eğitim Fakültesi, İlköğretim Matematik Öğretmenliği Programı 3. ve 4. sınıf öğrencileri ve aynı programdan mezun olmuş ilköğretim matematik öğretmenlerinin (1) değerlendirme okuryazarlık düzeylerini araştırmak ve (2) değerlendirme okuryazarlıklarındaki gelişimi etkileyen faktörleri belirlemeye çalışmaktır.

Bize Nasıl Yardımcı Olmanızı İsteyeceğiz?

Araştırmaya katılmayı kabul ederseniz, çalışmanın ilk kısmında sizden beklenen, Değerlendirme Okuryazarlığı Envanterini eksiksiz ve dikkatli bir şekilde cevaplamanızdır. Envanter ortalama 40 dakikada doldurulmaktadır. Çalışmanın ikinci kısmında, değerlendirme okuryazarlığı gelişimi etkileyen faktörleri belirlemek amacıyla sizlerle yaklaşık 40 dakika sürecek bir görüşme yapılacaktır. Görüşmeler sırasında daha sonra içerik analizi ile değerlendirilmek üzere cevaplarınızın ses kaydı alınacaktır.

Sizden Topladığımız Bilgileri Nasıl Kullanacağız?

Araştırmaya katılımınız tamamen gönüllülük temelinde olmalıdır. Çalışmanın

ikinci kısmında gönüllü olan öğretmen ve öğretmen adaylarıyla görüşmeler yapılacağı için çalışmanın ilk kısmında katılımcıların isim ve soy isim bilgileri gerekmektedir. Ancak cevaplarınız tamamıyla gizli tutulacak, sadece araştırmacılar tarafından değerlendirilecektir. Katılımcılardan elde edilecek bilgiler toplu halde değerlendirilecek ve bilimsel yayımlarda kullanılacaktır. Sağladığınız veriler gönüllü katılım formlarında toplanan kimlik bilgileri ile eşleştirilmeyecektir.

Katılımınızla ilgili bilmeniz gerekenler:

Çalışma, genel olarak kişisel rahatsızlık verecek sorular içermemektedir. Ancak, katılım sırasında sorulardan dolayı ya da herhangi başka bir nedenden ötürü kendinizi rahatsız hissederseniz cevaplama işini yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda çalışmayı uygulayan kişiye, çalışmadan çıkmak istediğinizi söylemek yeterli olacaktır.

Araştırmayla ilgili daha fazla bilgi almak isterseniz:

Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz. Çalışma hakkında daha fazla bilgi almak için ODTÜ öğretim üyelerinden Doç. Dr. Bülent Çetinkaya (E-posta: bcetinka@metu.edu.tr) ya da Beyza Şahin (E-posta: beyza.sahin@metu.edu.tr) ile iletişim kurabilirsiniz.

Yukarıdaki bilgileri okudum ve bu çalışmaya tamamen gönüllü olarak katılıyorum.

(Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

İsim - Soyad:

Tarih:

İmza:

E-mail:

Appendix F

METU HUMAN SUBJECTS ETHICS COMMITTEE APPROVAL

UYGULAMALI ETİK ARAŞTIRMA MERKEZİ
APPLIED ETHICS RESEARCH CENTER

ORTA DOĞU TEKNİK ÜNİVERSİTESİ
MIDDLE EAST TECHNICAL UNIVERSITY

DUMLUPINAR BULVARI 06800
ÇANKAYA ANKARA/TURKEY
T: +90 312 210 22 91
F: +90 312 210 79 59
ueam@metu.edu.tr
www.ueam.metu.edu.tr

20 HAZİRAN 2022

Konu: Değerlendirme Sonucu

Gönderen: ODTÜ İnsan Araştırmaları Etik Kurulu (İAEK)

İlgi: İnsan Araştırmaları Etik Kurulu Başvurusu

Sayın Doç.Dr. Bülent ÇETİNKAYA

Danışmanlığını yürüttüğünüz Beyza ŞAHİN'in "HİZMET ÖNCESİ VE HİZMET İÇİ ORTAOKUL MATEMATİK ÖĞRETMENLERİNİN DEĞERLENDİRME OKURYAZARLIĞININ İNCELENMESİ: ÖĞRETMENLERİN DEĞERLENDİRME OKURYAZARLIĞINI ETKİLEYEN NEDENLERİ ANLAMA" başlıklı araştırması İnsan Araştırmaları Etik Kurulu tarafından uygun görülerek gerekli onay **0337-ODTUİAEK-2022** protokol numarası ile onaylanmıştır.

Bilgilerinize saygılarımla sunarım.

Prof. Dr. Mine MISIRLISOY
Başkan

Doç. Dr. İ.Semih AKÇOMAK
Üye

Dr. Öğretim Üyesi Müge GÜNDÜZ
Üye

Dr. Öğretim Üyesi Şerife SEVİNÇ
Üye

Dr. Öğretim Üyesi Murat Perit ÇAKIR
Üye

Dr. Öğretim Üyesi Süreyya ÖZCAN KABASAKAL
Üye

Dr. Öğretim Üyesi A. Emre TURGUT
Üye