Professional Development Workshops: Determination and Fulfillment of Educational Requirements in Measurement-Evaluation

Gülden Gürsoy*, Esen Turan Özpolat

Adıyaman University, Adıyaman, Turkey

Corresponding author: Gülden Gürsoy, E-mail: gakdag@adiyaman.edu.tr

ARTICLE INFO

ABSTRACT

The present study aimed to determine the educational requirements for teachers in measurement-evaluation and whether the professional development (PD) workshops organized based on these requirements have an impact on evaluation literacy, attitudes, and self-efficacy perceptions of the teachers. The mixed-method design, where both quantitative and qualitative data were collected, was employed. A total of 204 middle school teachers participated in the study that aimed to determine the educational requirements of teachers in measurement-evaluation. For maximum diversity, 17 science teachers with different genders (10 male 7 female) and seniority between 3 and 15 years attended PD workshops. Five data collection instruments, including a survey form, three scales, and an interview form, were used to collect the study data. In quantitative analysis, descriptive statistics, dependent groups t-test, multilinear regression analysis were employed and the content analysis method was used in qualitative analysis. The analysis of the study data demonstrated that there were positive changes in assessment literacy, attitudes, and self-efficacy perceptions of the teachers towards measurement-evaluation, and it was determined that the workshop model was applicable in PD programs. As teachers' knowledge of assessment increased, they increasingly relied on self-knowledge and decided to adopt the new assessment techniques they learned. After the training requirements of the teachers were determined, it was recommended that the professional development training should be supported to meet their needs in workshops based on the study findings.

Key words: Professional Development, Workshops, Assessment Literacy, Attitudes, Self-Efficacy Perceptions

INTRODUCTION

Teacher training is the most basic process in the education system, it is extremely important for qualified education and instruction. Although teacher training has been taken seriously in every period, the requirements in teacher training have always changed with the changes in social conditions. According to Borko, Jacobs, and Koelner (2010), as a change was experienced from behaviorism to cognitivism and contingency, the ideas on cognitive learning and the nature of instruction have been developed. According to the authors, these changes also led to changes in teacher learning and professional development (PD) approaches. This change was dependent on variations in needs. The increasing interest in the PD of teachers in the 21st century was due to the fact that teacher training is considered an important method in the acquisition of complex skills required by the students in this period. Because, advanced instruction styles are required for students to develop complex problem-solving, critical thinking, self-management, effective communication, and cooperation skills (Darling, Hammond, Hyler &

Gardner, 2017). According to Patel (2007), teachers should not consider themselves as a knowledge resource but as an organizer of learning and learning experiences. Patel argued that the teacher is a carrier of culture, who structures the characters and personalities of the students. The teacher is the carrier of moral and ethical values, and a guide and consultant. The teacher is the architect of the future of the nation, inspiring the populace for social change. Thus, it is important to meet the training requirements of teachers, who undertake such significant tasks.

Adaptation of individuals to cultural, social, and economic developments due to scientific and technological advances and increasing knowledge is only possible through lifelong education (Yalın, 2001). The phenomenon of lifelong learning includes both students and teachers, as well as any employed or unemployed individual in any social class since it is suitable and necessary for individuals of all ages. In education and instruction, teachers are the primary factor in the lifelong learning process. Thus, PD of teachers should include training updates based on the requirements of lifelong learning.

In-service training programs focused on PD represent a form of lifelong education for the PD of individuals in the field of education, thereby providing educators an opportunity to expand and deepen their knowledge as well as a chance of training and renewal (Pepeler, Özbek, Adanır & Kılavuz, 2017). In-service training that could be provided throughout the professional careers of teachers entails a long and relatively more important education than the vocational teacher training (Patel, 2007). For this reason, it is recommended to remove the obstacles to their professional development, to provide lifelong learning opportunities, and to determine professional development education policies (Can, 2019). Furthermore, the content of career development training should focus on the development of teachers' pedagogical skills and student learning (Blank, de las Alas & Smith, 2007).

Vocational teacher training is a complex process that requires individual and collective, emotional and cognitive participation of the teachers (Avalos, 2011), especially after the studies that reveal that PD programs are high quality and effective in recent years (Antoniou & Kyriakides, 2013; Bolam, 1994; Buczynski & Hansen, 2010; Büyükşahin, 2018; Deglau & O'Sullivan, 2006; Doppelt, Schunn, Silk, Mehalik, Reynolds, & Ward, 2009; Giraldo, 2014; Greenleaf et al., 2011; Heller, Daehler, Wong, Shinohara & Miratrix, 2012; Jeanpierre, Oberhauser & Freeman, 2005; Kleickmann, Trobst, Jonen, Vehmeyer & Moller, 2016; Landry, Anthony, Swank & Monseque-Bailey, 2009; McMeeking, Orsi & Cobb, 2012; Meissel, Parr & Timperley, 2016; Önen, Mertoğlu, Saka & Gürdal, 2009; Randel, Apthorp, Beesley, Clark, & Wang, 2016; Resnick, 2005; Sathyanesan, 2001; Shaha & Ellsworth, 2013; Stegall, 2015) further PD opportunities are in urgent need, especially in recent years (Can, 2013; Borko et al., 2010). Can (2014) emphasized that teacher organization is insufficient, teachers' development cannot be achieved and there is a need for an organization that can provide professional development of teachers. It is necessary to eliminate the barriers to professional development training and to determine which professional development styles have the greatest impact on teacher practice and student achievement (Doppelt, Schunn, Silk, Mehalik, Reynolds & Ward, 2009).

The first step in the preparation of a good PD program is the needs analysis (Özdemir & Yalın, 2007). Academicians in the field of PD and teacher education agree that PD programs should respond to the needs of teachers and consider teachers as students of their teaching (Giraldo, 2014). Needs analysis in education helps in understanding the information that has been taught so far and is missing for later. Thus, it allows them to make informed decisions about what to do and to increase the scope and effectiveness of the training program (McCawley, 2009). It is seen that the programs prepared after the need determination gives positive results. Teachers generally tend to prefer training that suits their needs and supports their teaching processes (Demirel, 2009; Taymaz, 1997). In the study, the training needs of teachers for measurement-evaluation were determined and the content of the PD program was shaped according to the needs.

For the PD program to be efficient, care was taken to apply the six principles proposed by Supovitz and Turner (2000). These are: (1) It should be prepared with the model that occupies the participants with research, inquiry, and practices. (2) It should be given both teachers' intensive training and training should be continuous. (3) It should engage teachers with concrete teaching tasks and be based on teachers' experiences with students. (4) Focus on subject area knowledge and enrich teachers' content skills. (5) They should have specific standards and show how teachers should relate their work to improve their performance. (6) The reform strategies of the program should also be associated with other aspects of school change. In the study, the PD program has been programmed in such a way that teachers can actively participate, practice the theory, develop products that are unique to them through concrete tasks, and increase the pedagogical knowledge of teachers. In particular, the Workshop Model was used to transform teachers' learning in theory into practice. Töre (2017) not only makes learning permanent but also enables the spread of learned knowledge thanks to the cooperative learning of the PD programs supported by workshops.

PD programs can be prepared in many areas of education. PD programs; it arises due to the updating of existing situations, introducing and informing new approaches, or completing teachers' deficiencies in technology pedagogical knowledge. After the studies in the literature (Özenç & Çakır, 2015; Özbaşı & Çıkrıkçı Demirtaşlı, 2013; Akdağ & Ekmekçi, 2014), it is a measurement-evaluation in a field that should have PD training for teachers. The importance of measurement-evaluation for education has been asked by many educators and policymakers in the United States with the word "what gets assessed is what gets taught" (O'Day & Smith, 1993; Akt: Koh, 2011). In the last two decades, countries have started to pay attention to assessment and evaluation development programs for teachers to become literate teachers, after the understanding that teachers' in-class evaluation high performances provide students to show high-level skills (for example, complex thinking, reasoning, problem-solving, communication and conceptual understanding of the subject) (Smith & O'Day, 1990; Koh, 2011). The inability of teachers with not assessment literacy to provide reliable assessment results has revealed the necessity for each teacher to be assessment literacy after potential damage to their countries, schools, students, and families. However, studies conducted in the literature show that teachers' assessment literacy levels (Campbell, Murphy & Holt, 2002; Gürsoy & Aydoğdu, 2020; Mertler ve Campbell, 2005; Coombs, DeLuca, LaPointe-McEwan, & Chalas, 2018; Volante & Beckett, 2011; Lee & Son, 2015; Odo, 2016) and their attitudes are not at the desired level, revealed that teachers feel inadequate in assessment and evaluation (Arslan, İlker & Demirhan, 2013; Mutluer, 2015; Lee & Son, 2015; Reeves & Honig 2015; Mertler, 2009). Certain studies demonstrated that the assessment literacy of a teacher was associated with both self-efficacy perception and attitudes of the same. The presence of a correlation between knowledge and attitude (Quilter & Gallini, 2000), along with a positive

Professional Development Workshops: Determination and Fulfillment of Educational Requirements in Measurement-Evaluation

15

significant correlation between teachers' assessment literacy levels and their self-efficacy perceptions (Kruse, Impellizeri, Witherel & Sondergeld, 2020), revealed the significance of professional development program development. It was suggested that this topic should be researched to develop professional development program content which would improve both assessment literacy levels, attitudes, and self-efficacy perceptions of the teachers and to investigate the effectiveness of these. The present study aimed to determine the educational requirements of teachers in measurement and evaluation and whether the professional development workshops organized based on these requirements affect teachers' assessment literacy, attitudes, and self-efficacy perceptions.

METHOD

In the present study, the mixed method, where both quantitative and qualitative data were collected, was employed. In the mixed method, the researcher collects both quantitative (close-ended) and qualitative data (open-ended) to understand the research problems, integrates the two datasets, and then concludes with the advantages of the integration of these two datasets, and this research approach is employed in the fields of health, social and behavioral sciences (Creswell, 2017). In mixed studies, data diversification (triangulation) is recommended as a strategy to analyze the overall quality of data (Teddlie & Tashakkori, 2015). In the present study, a data diversification strategy was employed.

In this research, it is aimed to contribute to the clarity and comprehensibility of the data by supporting numerical data with qualitative data. While the screening method was employed to determine the workshop content in the quantitative dimension of the study, a single group pretest-posttest weak experimental design was adopted to determine the workshop achievements. In the qualitative dimension of the study, the interview technique was used to determine the views of the workshop participant teachers on education. The flowchart that reflects the mixed method employed in the study is presented in Figure 1.

Participants

In the present study, the first participant group was assigned with a simple random sampling method. According to Merriam (1998), purposive sampling provides the highest level of information for the case that the researcher desires to explore, understand, and comprehend in-depth. A total of 204 middle school teachers employed in Adıyaman Province central district in Turkey participated in the study and voluntarily completed the "educational requirements determination form" that was developed to determine the educational requirements of teachers in measurement-evaluation.

The second group was assigned with the maximum diversity sampling technique, a purposive sampling method, to collect in-depth data To obtain maximum diversity among 40 science teachers who desired to participate in the training announced by the Adıyaman Directorate of National Education after the development of the in-service training content, 20 female and male science teachers with 3-15 years seniority, who never attended professional development training were included in the study. The schools where these teachers were employed were also different. However, since three people could not attend the training regularly, the data were collected from 17 individuals. Participant demographics are presented in Table 1.

According to Table 1, ten participants were male and seven were female. The seniority of the participants varied between 3 and 15 years.

Data Collection Instruments

The employment of various data collection instruments in research improves validity and reliability (Vidovich, 2003). In the present study, five data collection instruments, including a survey form, three scales, and an interview form, were used to collect the quantitative and qualitative data. In the first section of the "educational requirements determination form (ERDF)" developed by the authors to

Table 1. Participant Demographics

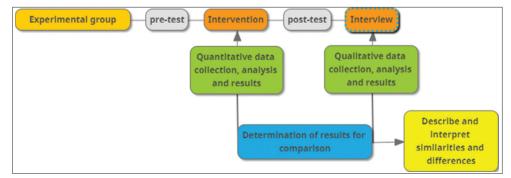| Code	Gender	Seniority	Code	Gender	Seniority
P1	M	15	P10	M	6
P2	M	10	P11	M	8
P3	M	12	P12	F	14
P4	M	4	P13	F	9
P5	M	4	P14	F	11
P6	F	11	P15	M	17
P7	F	3	P16	M	19
P8	M	5	P17	F	13
P9	F	6			

Figure 1. Mixed-Method Intervention Design (Creswell, 2017; 60)

determine the educational requirements of teachers in measurement-evaluation and completed by the participants on Google forms included items on complementary assessment techniques, the second section included items on traditional assessment techniques, the third section included assessment techniques conducted with web 2.0 tools, and the fourth section included items on the steps employed in achievement test development. The "*Assessment Literacy Inventory (ALI)*" (1993), developed by Plake and Impara in collaboration with American Education Assessment Association and W.K. Kellog Association in Nebraska-Lincoln University, translated to Turkish language by Gül (2011) including reliability and validity studies and aimed to determine the assessment literacy levels of the teachers was applied before the workshops. ALI includes seven core teacher competencies. There are five questions in each competence area. KR-20 Reliability coefficient of the literacy questionnaire was .718, Mean Difficulty was .556, and Mean Discrimination was 0.374.

Attitudes About Measurement-evaluation Scale (AMES) that was developed by Gül (2011) and included 4 factors was applied as pre-test and post-test to determine the attitudes of teachers towards measurement-evaluation. These factors included the sub-dimensions of special attention, anxiety-concern, importance, and competence. Measurement-evaluation Self-Efficacy Perception Scale (MESPS) developed by Kılınç (2011) was also used as a pre-test and post-test to determine teacher self-efficacy perceptions about measurement-evaluation. The overall Cronbach's alpha coefficient of the 23-item 5-point Likert-type scale was calculated as .96. The Cronbach's alpha coefficients for the first sub-dimension (knowledge) were .93 and the Cronbach's alpha coefficient for the second sub-dimension (skill) was .95.

Semi-structured interviews were conducted with the participants to determine the contribution of the workshop organized for the educational requirements of teachers in measurement-evaluation. The interview form developed by the authors included three questions. Since the interviews were conducted to support the quantitative data, the interview questions were developed based on the scale items. The questions aimed to determine whether there were differences in assessment literacy, self-efficacy perception levels of the teachers, and their attitudes towards measurement and evaluation. The interview form questions included the following:

1. In your opinion, what were the contributions of the workshops on measurement-evaluation?
2. What kind of differences have you observed in your interest or attitudes towards measurement-evaluation practices after the workshops?
3. Do you think there was a change in your competence in measurement-evaluation after the workshops?

Data Collection

Before training, the study content was developed and submitted for the approval of the Teacher Training Unit in Adıyaman Directorate of National Education. After the approval was obtained, the form developed to determine the teacher's educational requirements in measurement-evaluation was presented to the teachers via Google forms. The link to the form was distributed in WhatsApp groups and volunteers were asked to complete the form. Access to the form, which was available for about 15 days, was closed after the adequate number of feedbacks was received. After the form data was analyzed, the workshop program was developed based on teacher requirements. The development of the workshop content was conducted through the planning, development, implementation, and evaluation phases. In the first planning stage, the program content and methodology were determined.

I.   Planning
a.   Determination of the goals
     The main aim of the study was to improve the knowledge and skill levels of science teachers on measurement instruments in the context of traditional and complementary measurement-evaluation techniques, their assessment literacy levels, and their attitudes and self-efficacy towards measurement-evaluation.
b.   Determination of the workshop content
     A literature review was conducted to determine the workshop content and the content was determined based on teacher requests for measurement-evaluation. The details of the workshop content are presented in Table 2.
     When the Table 2 is examined, the names, goals and time zones of the professional development workshops are included.
c.   Determination of the method
     During the workshops, hands-on instruction was conducted with the constructivist approach. The instructors first completed the distributed worksheets in each activity individually, and then together in a group discussion. Thus, they restructured their knowledge in a social constructivist environment. Furthermore, an author led the group in each desk during the activities, while supporting the students during the reconstruction of the newly learned knowledge in the developmental process.
II. Development
     The workshops, which were designed as practical activities during the second week of the seminar (between 09: 00-12: 00), included presentations, visual demo activities, and group work supported by worksheets and inter-group discussions.
a.   Workshop schedule
     During the development of the workshops, each activity was selected individually, adapted, and the required time was determined and a schedule was developed in the study.
b.   Workshop sub-sections
1.   Presentation
     The authors conducted presentations to introduce the basic measurement-evaluation concepts, Web 2.0 tools, and to promote the workshop activities.
2.   Activities
     Groups were formed for each measurement tool, and the groups conducted activities.
1.   Table of specifications development
2.   Achievement test development
3.   Test and item analysis
4.   Vee and I-diagram plotting
5.   Development of material for student product files

Professional Development Workshops: Determination and Fulfillment of Educational Requirements in Measurement-Evaluation

17

Table 2. Workshop Program

Date	Duration	Topic	Workshops
18.06.2018 (1st Day)	09.00-10.30	Introduction of the program, pretest	"Games to meet: Drama"
	11.00-13.00	What is measurement-evaluation? What are traditional and complementary measurement-evaluation techniques? What is the significance of measurement-evaluation in the general structure of the 2018 Science Curriculum?	"Find the differences" workshop
19.06.2018 (2nd Day)	09.00-11.00	What is the student product file? What is the goal of the student product file? What is the content of student product files and how student product files are evaluated?	"Designing my product file" workshop
	11.30-13.00	What is the significance of mind maps as a measurement-evaluation tool in science instruction? What are the priorities in the application and evaluation of mind maps?	"I measure with maps" workshop
20.06.2018 (3rd Day)	09.00-11.00	What is the significance of word association, prediction-observation-explanation, concept cartoons as measurement-evaluation tools in science instruction? What is their intended use? What are the priorities in their application and evaluation?	"I get to know contemporary measurement techniques" workshop
	11.30-13.00	What is the significance of posters, Vee diagrams, and projects as measurement-evaluation tools in science instruction? What is their intended use? What are the priorities in their application and evaluation?	"I get to know new measurement tools" workshop
21.06.2018 (4th Day)	09.00-13.00	What is experience-based open-ended question development? What are the priorities in open-ended question development? How to score open-ended questions? How to ensure scoring reliability?	"I develop questions with stories" workshop
22.06.2018 (5th Day)	09.00-11.00	What are the achievement test development steps? What are the priorities in a reliable and valid achievement test development? What is the significance of the analysis of test results? How to conduct item and test analysis? How to score the test?	"I develop my own tests" workshop
	11.30-13.00	Which Web 2.0 tools are employed in measurement-evaluation?	"Gamified assessment" workshop
	13.00 13.30	Post-test	

6. Poster design
7. Mind map, word association activities
8. Graded score key development
9. Gamified measurement-evaluation

III. Application

The measurement tools developed in the activity were applied by each teacher in their classrooms, and methods were selected to interpret the collected data and to communicate with the stakeholders.

After the draft workshop program was developed, the activities associated with the workshop content, worksheets, sample portfolios, topic summaries for the force-motion unit, examples of daily life questions, graded scoring key, mind maps, etc. were turned into a booklet and published for distribution to the participants. Furthermore, five round tables for four people were prepared and colored pens, paper sheets, scissors, and glue were placed on each table. During the workshops, the content was summarized using topical examples before each activity, and applications were initiated by emphasizing the priorities. The authors engaged with each group during the activities and provided verbal feedback to allow them to comprehend the work better. The products developed by the teachers were collected every day and more detailed feedback was provided the day after. After the in-service training program, the pretest scales were reapplied as posttest and 20-minute interviews were held with the teachers immediately. After the training program, a WhatsApp group was created to maintain the communications, and the participants were told to communicate with the authors for the solution to their problems they could experience during the applications. The study flowchart is presented in Figure 2.

IV. Evaluation

The scales were applied before the workshops as posttest and semi-structured interviews were conducted with the teachers. Furthermore, Adıyaman National Education Directorate R&D unit staff applied the in-service training evaluation form. The training evaluation findings were reported to the relevant unit in the Ministry of National Education.

Data Analysis

Data analysis in mixed-method studies includes the processes where both qualitative and quantitative data are collected, analyzed, correlated, and combined (Teddlie & Tashakkori, 2015). After the quantitative data collected with the scales were coded
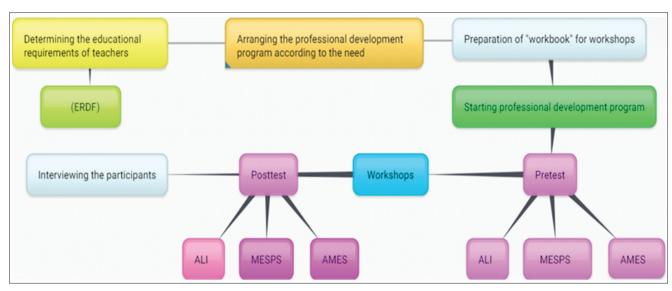
Figure 2. Study Flow-Chart

Tablo 3. "ALI, MESPS, AMES" Shapiro-Wilks Normality Analysis Results

Test	Shapiro-Wilks	p	Skewness	Kurtosis
ALI-Pretest	.886	.019	-.439	-.1.13
ALI-Posttest	.928	.119	-.472	-1.064
MESPS-Pretest	.967	.761	-.049	-.313
MESPS-Posttest	.905	.084	.503	-1.148
AMES-Pretest	.880	.032	-1.393	1.588
AMES-Posttest	.967	.761	.-594	.019

into a data analysis software, the authors decided which tests should be conducted to serve the study aim. Normality assumptions should determine which parametric and nonparametric tests should be used in analysis instead of the participant size (Brewer, 1985). In the study, the tests were determined based on the normal distribution of the data. In the study, Shapiro-Wilks (SW) and skewness-kurtosis coefficient values were used to determine whether the data exhibited normal distribution. Shapiro- Wilks is used when the group size is less than 50 and Kolmogorov-Smirnov (KS) is used when the group size is larger than 50 (Büyüköztürk, Çakmak, Akgün, Karadeniz & Demirel, 2009). Furthermore, descriptive statistics were reported with mean and standard deviation. The skewness and kurtosis coefficients and Shapiro-Wilk normality analysis results that were conducted to determine the statistical Inventory (ALI), Measurement-evaluation Self-Efficacy Perception Scale (MESPS) and Attitudes towards Measurement-evaluation Scale (AMES) are presented in Table 3.

The normal distribution of the student ALI, MESPS, AMES scores were determined with the Shapiro-Wilks test and skewness-kurtosis coefficients. If the p-value in the Shapiro-Wilks test is greater than .05, the data distribution is normal. Since the skewness and kurtosis coefficients for the data were between -1.96 and +1.96, the distribution of the data was considered normal. Thus, parametric tests were used in the study. The quantitative questions employed in the study and associated statistical processes are presented in Table 4.

The records of the interviews conducted with the teachers who participated in the workshops constituted the qualitative study data. In qualitative analysis, the content analysis method was employed. The content analysis method was used in qualitative analysis in the study. In qualitative data analysis, the coding technique was used. In coding technique, the researcher organizes the collected data to make it more understandable, conceptualizes and categorizes the concepts based on their associations, and establishes connections between the categories, interprets, and reports these connections (Strauss and Corbin, 1990). As in the conceptualization phase, the categories are named. However, the names assigned to the categories are more abstract when compared to the names assigned to the concepts. During the interviews, three questions were asked to the teachers, and the data collected in the interviews were grouped under 17 codes and 4 categories and presented in a single table.

Validity and Reliability

Validity in qualitative research means that researchers reflect and report their studies objectively and with all possible details (Kirk & Miller, 1986 as cited in Yıldırım & Şimşek, 2011). To ensure internal validity (credibility) in the study, it was ensured that the research questions and data were approved by an expert in the qualitative field and a professor in the field of science education. Also, to ensure internal validity, participant confirmation was obtained in terms of the comprehensibility of the research questions and the consistency of the researcher and participant comments on the data obtained. After the analysis of the data to ensure internal reliability, some findings are presented by quoting directly without comment. The fact that the interviews were conducted face to face also benefited the validity of the research. The research has been reported in all its details in terms of determining the participants, data collection, and data analysis processes. Paying attention to the consistency of the research questions with the subject and the richness of sample selection was deemed important to ensure external validity in the study.

Professional Development Workshops: Determination and Fulfillment of Educational Requirements in Measurement-Evaluation

19

Table 4. Quantitative Data Analysis

Research Question	Data collection instrument	Data analysis
1. What are the requirements of the teachers in measurement-evaluation?	ERDF	Descriptive Statistics
2. Is there a significant difference between assessment literacy pretest and posttest scores of science teachers who participated in the PD workshops?	ALI	Dependent Groups t-test
3. Is there a significant difference between measurement-evaluation self-efficacy perception pretest and posttest scores of science teachers who participated in the PD workshops?	MESPS	Dependent Groups t-test
4. Is there a significant difference between measurement-evaluation attitude pretest and posttest scores of science teachers who participated in the PD workshops?	AMES	Dependent Groups t-test
5. Do measurement-evaluation self-efficacy perceptions and measurement-evaluation attitudes of science teachers who participated in the PD workshops predict their assessment literacy?	ALI MESPS	Multilinear Regression Analysis

The data obtained for the reliability of the study were examined separately by both researchers, discussed within the framework of consensus and disagreement, and the final version was reached with the arrangements made. The reliability formula of Miles and Huberman (1994, p. 64) was used to calculate the reliability of the interview data:

$$\text{Reliability} = \text{Agreement}/(\text{Agreement} + \text{Disagreement}) \times 100$$

As a result of the calculation, a reliability number of 84.00% has been reached, which indicates that this research is reliable. Because Miles and Huberman (1994) state that the value obtained according to the formula for reliability should be 70% of the minimum score.

The measurement reliability of the scales used to collect quantitative data in the study was calculated separately. In this study, it was deemed appropriate to use the α coefficient, which is considered to be the internal consistency criterion and developed by Cronbach, to calculate the reliability level of the MESPS, a Likert-type scale. Measurement reliability was .77 for AMES, and the measurement reliability was. It was determined as 86. The KR-20 coefficient was calculated as the internal consistency criterion of the CALI coded as 1 and 0 and was determined as .71.

FINDINGS

The findings on the quantitative and qualitative study data are presented in two separate sections; however, to obtain reliable and valid results, the two data types were integrated into the discussion section.

Quantitative Findings

The quantitative study findings are presented in different sections based on research sub-problems.

Findings on Teachers' Educational Requirements in Measurement-Evaluation

The descriptive statistics of the data collected with the ranking of the preferences of 204 teachers employed in Adıyaman about the topics they required training the most are presented in Table 5.

The review of Table 5 demonstrated that the teachers predominantly desired training on student product files (65) among the complimentary measurement-evaluation techniques. The student product file was followed by worksheets (39), mind maps (38), word association (36), prediction-observation-explanation (34), project (33), conceptual cartoons (33), and graded scoring key (30). It was determined that the teachers desired training on a structured grid (22) and descriptive branched tree (21) techniques the least. It was observed that teachers mostly desired training on open-ended question development among the traditional measurement-evaluation techniques. The open-ended question development (77) was followed by multiple-choice (57), matching question (35), filling the blanks (34), and true-false (33) question development techniques. It was determined that they desired training on measurement-evaluation with Quizz application (87). The Quizzz tool was followed by Kahoot (54) and Quizlet (28). The analysis of the training required for the achievement test development steps demonstrated that the teachers mostly desired training on the table of specifications development. This (66) was followed by test development (55), question content development (49), item-test analyses (44), and conversion to standard score (33).

Findings on the Impact of PD Workshops on Assessment Literacy of Science Teachers

Descriptive data and dependent groups t-test results on the "Is there a significant difference between assessment literacy pretest and posttest scores of science teachers who participated in the PD Workshops?" sub-problem are presented in Table 6.

Table 6 demonstrated that the assessment literacy levels of the teachers changed significantly after the workshops. The difference between the mean assessment literacy pretest score ($\bar{x} = .56$) and mean post-test score ($\bar{x} = .66$) of the teachers favored the mean post-test score ($\bar{x}$ posttest $> \bar{x}$ pretest) and dependent groups t-test results showed that the difference was significant ($t = -5.00$; $p < .05$). It could be suggested that the workshop-assisted in-service

Table 5. The Topics that Teachers Required Training in Measurement-Evaluation

Category	Code	f
Complimentary Assessment Techniques	Student Product File	65
	Worksheets	39
	Mind Maps	38
	Word Association	36
	Prediction-Observation-Explanation	34
	Project	33
	Conceptual Cartoons	33
	Graded Scoring Key	30
	Grid	22
	Descriptive Branched Tree	21
Traditional Assessment Techniques	Open-ended Question Development	77
	Multiple Choice Question Development	57
	Matching Question Development	35
	Filling the blanks Question Development	34
	True/False Test Development	33
Web 2.0 Tools	Quizz	87
	Kahoot	54
	Quizlet	28
Test Development	Table of Specifications Development	66
	Test Development	55
	Question Content Development	49
	Item-Test Analyses	44
	Conversion to Standard Score	33

Table 6. The dependent groups t-test results on the comparison of ALI Pretest and Posttest scores

Test	n	$\bar{x}$	SD	t	p
Pretest	17	.56	.085	-5.00	.000
Posttest	17	.66	.100		

Table 7. The dependent groups t-test results on the comparison of MESPS Pretest and Posttest scores

Test	n	$\bar{x}$	SD	t	p
Pretest	17	3.46	.492	4.441	.000
Posttest	17	4.05	.498		

measurement-evaluation training program could improve the assessment literacy levels of the teachers.

Findings on the Impact of PD Workshops on Measurement-Evaluation Self-Efficacy Perceptions of Science Teachers

Descriptive data and dependent groups t-test results on the "Is there a significant difference between measurement-evaluation self-efficacy perception pretest and posttest scores of science teachers who participated in PD workshops?" sub-problem are presented in Table 7.

Table 7 demonstrated that the measurement-evaluation self-efficacy perceptions of the teachers changed significantly after the workshops. The difference between the mean measurement-evaluation self-efficacy perception mean pretest score ($\bar{x}$ = 3.46) and mean posttest score ($\bar{x}$ = 4.05) favored the posttest score ($\bar{x}$ posttest> $\bar{x}$ pretest) and dependent groups t-test results showed that the difference was significant (t = 4,441; p <.05). It could be suggested that the workshop-assisted in-service measurement-evaluation training

program could improve the measurement-evaluation self-efficacy perceptions of the teachers.

Findings on the Impact of PD Workshops on the Attitudes towards Measurement-Evaluation of Science Teachers

Descriptive data and dependent groups t-test results on the "Is there a significant difference between measurement-evaluation attitude pretest and posttest scores of science teachers who participated in the PD workshops?" sub-problem are presented in Table 8.

Table 8 demonstrated that the attitudes of the teachers towards measurement-evaluation changed significantly after the workshops. The difference between the mean pretest teacher attitudes towards measurement-evaluation score ($\bar{x}$ = 3.22) and mean posttest score ($\bar{x}$ = 3.36) favored the posttest score ($\bar{x}$ posttest > $\bar{x}$ pretest) and dependent groups t-test results showed that the difference was significant (t = 4.015; p < .05). It could be suggested that the workshop-assisted in-service measurement-evaluation training

Professional Development Workshops: Determination and Fulfillment of Educational Requirements in Measurement-Evaluation

21

program could improve the attitudes of the teachers towards measurement-evaluation.

Findings on the Prediction of Assessment literacy by Measurement-evaluation Self-Efficacy Perception and Attitude Scores

Descriptive data and multiple regression analysis results on the "Do measurement-evaluation self-efficacy perceptions and measurement-evaluation attitudes of science teachers who participated in the PD workshops predict their assessment literacy?" sub-problem are presented in Table 9.

The results of the multilinear regression analysis conducted to determine whether variables such as self-efficacy perceptions and attitudes, which were considered to have an impact on assessment literacy, predicted their literacy levels demonstrated that there were significant correlations between these variables and literacy levels ($R = .594$; $R^2 = .352$) ($F_{(2-14)} = 3.807$ p <.05). These two variables explained 35% of the variance in literacy levels. Based on the standardized regression coefficients, the relative significance order of the predictor variables on literacy was self-efficacy ($\beta=.551$), and attitude ($\beta=.300$). Based on the significance tests conducted on the regression coefficients, it was observed that only the self-efficacy (p <.05) variable was a significant predictor of literacy. The analysis of the correlations between predictor variables and literacy revealed a correlation between literacy and self-efficacy ($r = .514$) and attitude ($r = .230$). Based on the results of the regression analysis, the regression equation to predict the literacy level was as follows:

$$\text{Literacy} = 18.025 + (-3.784 \times \text{Self-Efficacy Perception Score}) + (5.814 \times \text{Attitude Score})$$

Qualitative Findings

Qualitative findings and the interpretations based on these findings are included in this section of the study. The views of the teachers who participated in the workshops on an applied in-service training program were grouped under four categories. These codes and related frequencies are presented in Table 10.

The awareness category included the teacher statements on their awareness about their inner worlds. During the in-service training, teachers stated that realized the significance of measurement-evaluation (f = 7), that measurement-evaluation was not difficult (f = 5), the significance of process assessment (f = 3), the need to follow up technology (f = 3), and learning were possible during measurement (f = 2), respectively. Awareness is a cognitive and physical practice that includes focusing attention on current experiences and observing inner experiences (Kabat-Zinn, 2005). Conscious awareness is awakening. This entails directing our attention and awareness to our current experiences with all our heart and acceptance, which requires an intent to live the moment fully (Germer, Siegel, & Fulton, 2005). This awakening experienced by teachers about measurement-evaluation was an important step towards change and development. The fact that the activities conducted in the process had such an effect was desirable progress in the study. Certain direct quotes of the teachers are as follows in this category:

P₂: I realized that I should assess the process more actively with tools such as vee diagrams and mind maps. I learned to use these techniques at the middle school level.

P₇: Thanks to these workshops, my awareness of measurement-evaluation rose and this awareness led to motivation. I think that I can employ alternative measurement more effectively in the classroom after the training.

In the category of change, the decisions of the teachers that they would reflect the training in classroom education were observed. Teachers decided that they will change their measurement habits (f = 8), increase the number of measurement tools they employ in the classroom (f = 6) and change their measurement criteria (f = 3). The fact that teachers used expressions stating that they have made decisions on in-classroom measurement-evaluation activities during the interviews may indicate that they were willing to implement the new measurement tools they learned in future problems. Certain direct quotes of the teachers are as follows in this category:

P₅: I decided to change the measurement-evaluation techniques I implemented before and to add new criteria.

P₈: The workshops were very effective. For sure, it led to a serious awareness of what we do and do not. Now, I think I need to change the techniques we use in the class.

In the development category, it was determined that the teachers included statements about the changes they experienced after the in-service training. It was determined that the teachers started to feel more competent (f = 9) about measurement-evaluation due to the practices they conducted during the training. The development of teacher competences was observed more specifically in complementary

Table 8. The dependent groups t-test results on the comparison of AMES Pretest and Posttest scores

Test	n	x̄	SD	t	p
Pretest	17	3.22	.176	4.015	.001
Posttest	17	3.36	.208		

Table 9. Multiple Regression Analysis Results on the Predictive Power of Assessment Literacy

Variable	B	Standard Error	β	t	p	Paired r	Partial r
Constant	18.025	14.659	-	1.230	.239	-	-
Self-Efficacy	-3.784	1.488	-.551	-2.543	.023	-.514	-.562
Attitude	5.814	4.208	.300	1.382	.189	.230	.346

R=.594, R²=.352, Corrected R²=.260, $F_{(2-14)}$=3.807, p=.048

Table 10. The Views of Teachers about PD Workshops

Category	Code	f
Awareness	Recognizing the significance of measurement-evaluation	7
	Recognizing that measurement-evaluation is not difficult.	5
	Recognizing the significance of process assessment.	3
	Recognizing the need to follow up the technology	3
	Recognizing the fact that students could learn during the evaluation	2
Change	Deciding to change measurement habits	8
	Deciding to increase the number of measurement tools utilized in the classroom	6
	Deciding to change the criteria	3
Development	Feeling more competent after the intervention	9
	Elimination of deficiencies in complementary measurement-evaluation	6
	Learning about Web 2.0 tools employed in the measurement	4
	Learning how to implement the theoretically recognized tools	4
	Starting to communicate with colleagues in the same province	3
Affective	Increase in the belief that it could be implemented as one learns about it	9
	Increase in the interest and desire to implement more current approaches	6
	Increase in motivation after the workshops	4
	Recognizing that implementation of new techniques is exciting	3

measurement-evaluation (f = 6), Web 2.0 tools (f = 4), and application of certain techniques (f = 4), respectively. During the training, the fact that the complementary measurement-evaluation techniques were introduced and the practices they conducted could be a factor in the improvement of their knowledge. Furthermore, the efforts of teachers to produce products together during the workshops allowed them to communicate more with each other (f = 3) and to get to know one another. The workshop model is an education method that includes practical activities rather than traditional instruction and is based on teamwork and active communication (Lumpe, 2007; Gosser et al., 1998). It could be suggested that the workshop model (Russell, Hey, Thoen & Walz, 1978), which is employed to acquire new skills and allow interaction, served its purpose based on teacher statements. Certain direct quotes of the teachers are as follows in this category:

> P_{17}: My knowledge and perception of the measurement improved since we conducted hands-on studies in the

workshop. I think that I can now implement measurement-evaluation better.

> P_{14}: I learned to use new web tools. I also observed that active measurement-evaluation is based on process assessment.

In the affective category, it was determined that the teachers included statements that expressed the affective changes they experienced after the in-service training. It was determined that the teachers learned in more detail since they conducted practical activities during the training, and their beliefs in the implementation of that knowledge increased (f = 9). The other statements in this category reflected the increase in teachers' desire and interest in implementing more current approaches (f = 6), motivation to implement these approaches in the classroom (f = 4), and considering the process exciting (f = 3). The fact that teachers experienced such positive affective changes after the training may indicate that short-term training is effective. Certain direct quotes of the teachers are as follows in this category:

> P_{10}: I was already interested in measurement-evaluation applications. My reason for participating in the training also stemmed from this interest. After the training, my interest increased. I am also motivated when I think about the applications that I remember and recently learned that could contribute to in-class applications.

> P_{16}: My perspective on measurement-evaluation applications changed a lot. Until now, I used mostly traditional measurements and techniques, but I will use alternative measurement-evaluation techniques both during learning and measurement. Thanks to this training, my knowledge was refreshed and my interest in measurement increased.

DISCUSSION AND CONCLUSION

In the PD program organized as a "development training", the educational requirements of teachers in measurement-evaluation were initially determined. It was not recommended to move on to further steps without fulfilling the necessary conditions for this first step. Because, determining the required education would allow the interest in the program, and lead to a successful and efficient training (Bolam, 1994). Furthermore, the teachers are desired to implement the knowledge they learn in training, the in-service training program content should attract their attention and meet their needs (O'Sullivan, 2000). In the study, the educational requirements of teachers were collected using a questionnaire. The 204 middle school teachers who participated in the survey reported that they desired the training to include student product files (portfolios) among the complementary measurement-evaluation tools, open-ended question development among the traditional measurement-evaluation tools, Quizizz, among the web 2.0 tools used for measurement-evaluation, and table of specifications among the test development steps. In a study by Metin and Özmen (2010) that aimed to determine the in-service training requirements of science and technology teachers in performance evaluation, it was determined that teachers required training on performance evaluation and performance evaluation tools such

Professional Development Workshops: Determination and Fulfillment of Educational Requirements
in Measurement-Evaluation

23

as anecdote records, checklists, graded scale, graded scoring key, product files, peer and self-evaluation forms.

It was determined that there was a significant difference between the mean assessment literacy pretest and posttest scores of the teachers who participated in the PD. Similarly, the "development" category determined with the analysis of the qualitative study data demonstrated that the knowledge of the teachers on complementary measurement-evaluation improved, they learned to conduct measurement-evaluation with Web 2.0 tools, and learned to implement the techniques that they know in theory but could not apply, supporting the quantitative study data. Çoruhlu, Er Nas, and Çepni (2008) reported similar findings in their study. After the measurement and evaluation in-service training program, they reported that the teacher skills on complementary measurement-evaluation techniques such as student product files, performance evaluation, descriptive branched trees, and structured grids.

The number of activities conducted by teachers with higher assessment literacy levels increased (Mertler & Campbell, 2005), and they could communicate better with students and their parents since they could interpret student achievements more accurately (Brookhart, 1999). Siegel and Wissehr (2011) reported that teachers should understand the theoretical principles associated with measurement-evaluation and employ evaluation methods in the classroom to be considered as measurement-evaluation literate. They even argued that theory and practice should intersect and interact. Thus, the in-service training was not limited to workshops but also instructions on the significance of measurement-evaluation, theories, and evaluation methods in measurement-evaluation were provided for the teachers. Allen and Flippo (2002) argued that the introduction of measurement-evaluation methods on broad and different platforms would allow teachers to have in-depth knowledge of the measurement-evaluation approaches. Thus, the information about measurement-evaluation was presented to the teachers using the workshop model, different from the methods used in previous in-service programs. The fact that the participating teachers in the workshop-supported in-service program improved their assessment literacy levels demonstrated that this model could be used in future in-service training programs. Stegall (2015) reported that the training supported by workshops and presentations improved the quality of teacher instruction in the classroom and stated that it helps the teachers to realize their development and the improvements in student achievements. In an experimental study, Doppelt et al. (2009) determined that the students of the teachers who participated in a workshop-assisted PD program were more successful when compared to the students of the teachers who did not participate in the program. Furthermore, Doppelt et al. (2009) stated that they organized the workshops not only to develop the content or pedagogical knowledge of the teachers but also to create a community of learners among the teachers. Similar qualitative findings were reported in the present study. The teachers stated that they had the opportunity to "improve communication with their colleagues in the province" after the workshops in the present study. Workshops often lead to a higher level of interaction between the presenter and participants and include activity-based learning for the teacher participants.

It was determined that there was a significant difference between the mean measurement-evaluation self-efficacy pretest and post-test scores of the teachers who participated in the PD workshops. The qualitative study data supported the quantitative findings in the present study. Teacher statements that they felt more competent after the intervention supported the quantitative data which reflected the increase in the measurement-evaluation self-efficacy of the teachers. The codes "elimination of the deficiencies in complementary measurement-evaluation techniques" and "learning how to use the tools recognized theoretically" determined in post-PD program interviews demonstrated an improvement in teacher knowledge levels. The increase in teacher knowledge levels made them feel more competent, facilitating better measurement practices in the classroom. Because Tabarlet (1994) reported that teacher knowledge on measurement-evaluation was an important factor that affected the implementation of complementary measurement-evaluation tools in classrooms by the teachers. Corcoran, Dershimer, and Tichenor (2004) gradually investigated the knowledge and views of teachers on alternative evaluation techniques in their study. They instructed the teachers on how to use rubrics, portfolios, and checklists. They reported that when the self-confidence of the teachers improved, they started using other techniques than previously utilized alternative assessment techniques voluntarily. In the third stage, teachers became competent in knowledge and applications about alternative techniques. Şanlı and Pınar (2017) investigated the effects of a complementary measurement-evaluation development program on pre-service teachers' measurement-evaluation self-efficacy perceptions. Similarly, they concluded that pre-service teachers' measurement-evaluation self-efficacy perceptions improved in the selection, implementation, and evaluation dimensions of measurement-evaluation self-efficacy.

It was determined that there was a significant difference between the mean attitude towards measurement-evaluation pre-test and post-test scores of the teachers who participated in the workshop-assisted in-service training program. The analysis of the interview data revealed that the obtained qualitative codes supported quantitative study data. The codes "increased belief in applicability with more knowledge", "increased desire and interest in the implementation of more current approaches", "increased motivation due to the workshops," and the fact that they found "implementation of the new techniques exciting" demonstrated that teachers experienced affective awareness or changes during the workshops. The emergence of these affective changes within a short time may be due to the organization of the workshop content based on teacher requirements in the study. Süleymangil (2013) reported that the development and implementation workshops organized for teachers led to successful outcomes, the communication skills and motivation of the teachers with their colleagues improved, as well as their professional skills. Çam and Üstün (2016) reported that as professional attitudes increase, life-long learning tendency increases, and as the professional attitudes decrease, life-long learning tendency decreases. The teachers' continuous desire to learn

and implement measurement-evaluation techniques in the classroom after related PD programs would only be possible if they develop positive attitudes.

Furthermore, the interview findings demonstrated that teachers had a certain level of awareness. The codes about the significance and easiness of measurement-evaluation, and those where teachers stated that learning was possible during measurement and it was imperative to follow the technology were included in the awareness category. In the study, teachers' awareness of certain situations were factors that supported the first steps towards personal changes. Awareness is an important factor that strengthens the relationship between attitude and behavior. Because high awareness leads to a better understanding of attitudes and attitudes are more easily called into memory. Furthermore, when we generally behave, we pay attention to our attitudes towards that situation and allow this attitude to guide the behavior (Kağıtçıbaşı, 2006).

The interview findings also demonstrated that teachers made new decisions about themselves, and these were grouped under the category of "change". It was observed that teachers decided to change the measurement tools used in their classrooms, increase the number of measurement tools, and change the criteria they utilized in measurement. The decision is a mental process where one alternative deliberately selected to achieve the desired outcome (Oğuz, 2009). In other words, decision making is a fight against problems and uncertainties to eliminate them and to reveal what, how, and when (Bies & Moag, 1986). It was determined in previous studies that teachers preferred traditional measurement-evaluation techniques due to their habits (Gelbal & Kelecioğlu, 2007; Watt, 2005). Thus, the fact that the teachers who participated in the professional development program decided to change their habits indicated that they would increasingly prefer contemporary approaches instead of traditional ones.

It was determined that there were correlations between measurement-evaluation self-efficacy perceptions of the teachers who participated in the PD workshops and their attitudes towards measurement-evaluation, and assessment literacy levels, the measurement-evaluation self-efficacy perceptions of the teachers and their attitudes explained 35.2% of the variation in their assessment literacy. It was concluded that the developed regression model could be used to predict the assessment literacy of the teachers based on their measurement-evaluation self-efficacy and attitudes. In their study, Quilter and Chester (1998) stated that teachers develop a more positive attitude towards assessment and evaluation as assessment literacy improves.

After the teachers participated in the PD workshops, as a result of the analysis of the quantitative and qualitative data, it was determined that
1. Creating products in collaboration increases communication with colleagues.
2. Teachers had the opportunity to apply what they knew or heard in theory, so they became aware of themselves about what they know and what they don't.
3. With the increasing knowledge of assessment, teachers have started to rely on themselves and they made decisions that they will use the new assessment techniques they learned in their classrooms.

4. With the reference to the positive change of teachers' assessment literacy, attitudes, and their self-efficacy perceptions towards measurement-evaluation, it was determined that the workshop model was applicable in PD programs.

These findings indicated that PD workshops in the field may be adequate for various branches. Based on the study findings, the following are recommended:
1. For the professional development of teachers, their educational requirements should be regularly screened.
2. Development programs should be developed based on educational requirements in an environment where teachers could actively participate and practice.
3. Both quantitative and qualitative data on the effectiveness of professional development programs should be collected, the associated problems should be determined and the program should be reviewed.
4. The professional development program content should be developed to support cognitive, affective and psychomotor skills of teachers.
5. The development of professional development programs should be conducted by specialists in more specific fields and should be included in official educational policies.

REFERENCES

Akdağ, G., & Ekmekçi, S. (2014). Fen ve teknoloji öğretmenlerinin ölçme değerlendirmeye ilişkin yeterlilik algıları ve görüşleri (Adıyaman ili örneği). *Route Educational and Social Science Journal, 2*(3), 253-273.

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

Antoniou, P., & Kyriakides, L. (2013). A dynamic integrated approach to teacher Professional development: Impact and sustainability of the effects on improving teacher behaviour and student outcomes. *Teaching and Teacher Education, 29*, 1-12.

Arslan, Y. T., İlker, G. T., & Demirhan, G. T. (2013). Geliştirilen bir ölçme ve değerlendirme gelişim programının beden eğitimi öğretmen adaylarının ölçme ve değerlendirmeye ilişkin algılarına etkisi. *Kuram ve Uygulamada Eğitim Bilimleri, 13*(2), 1107-1124.

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

Bies, R. J., & Moag, J F. (1986). Interactional justice: Communication criteria of fairness. In R. J. Lewicki, B. H. Sheppard, & M. H. Bazerman (Eds.), *Research on negotiations in organizations* (pp. 43-55). Greenwich, CT: JAI Press.

Blank, R. K., de las Alas, N., & Smith, C. (2007). *Analysis of the quality of professional development programs for mathematics and science teachers.* Washington, DC: Council of Chief State School Officers.

Bolam, R. (1994). The impact of research on policy and practice in continuing professional development. *British Journal of ln-service Education, 20*(1), 35-44.

Professional Development Workshops: Determination and Fulfillment of Educational Requirements in Measurement-Evaluation

25

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2013). *Meta Analize Giriş* (S. Dinçer, Trans.). Ankara: Anı Yayıncılık.

Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. In Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education 7*, (pp. 548-556). Oxford: Elsevier.

Brewer, J. K. (1985). Behavioral statistics textbooks: Source of myths and misconceptions. *Journal of Educational Statistics, 10*(3), 252-268.

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

Buczynski, S., & Hansen, C. B. (2010). Impact of professional development on teacher practice: Uncovering connections. *Teaching and Teacher Education, 26*(3), 599-607.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2009). Bilimsel araştırma yöntemleri. Ankara: Pegem Akademi.

Büyükşahin, Y. (2018). *Araştırmaya dayalı fen öğretimine yönelik hazırlanan mesleki gelişim programı boyunca sınıf öğretmenlerinin gelişimlerinin izlenmesi* (Unpublished doctorate thesis). Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Çam, E. & Üstün, A. (2016). Öğretmenlerin mesleki tutumları ile yaşam boyu öğrenme eğilimleri arasındaki ilişkisi. *Hitit Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 9*(2), 459-476.

Campbell, C., Murphy, J. A., & Holt, J. K. (2002, October). *Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers.* Paper presented at the MidWestern Educational Research Association, Columbus, OH.

Can, E. (2013). Öğretmenlik kariyer sistemi: Sorunlar ve öneriler. IV. Eğitim Yönetimi Forumu (EYFOR-IV) in (s.71). 3-5 Ekim 2013, Balıkesir.

Can, E. (2014). Öğretmenlerin meslekî örgütlenme sorunları ve öneriler. V. Eğitim Yönetimi Forumu (EYFOR-V) in (s.152-153). Necmettin Erbakan Üniversitesi, 11-13 Eylül 2014, Konya.

Can, E. (2019). Öğretmenlerin meslekî gelişimleri: Engeller ve öneriler. *Eğitimde Nitel Araştırmalar Dergisi – Journal Of Qualitative Research In Education, 7*(4), 1618-1650. doi: 10.14689/Issn.2148-2624.1.7c.4s.14m

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrance Earlbaum Associates.

Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education, 71*, 134-144.

Corcoran, C. A., Dershimer, E. L., & Tichenor, M. S. (2004). A teacher's guide to alternative assessment taking the first steps. *The Clearing House, 77*(5), 213-216.

Çoruhlu, T. Ş., Er Nas, S., & Çepni, S. (2009). Fen ve teknoloji öğretmenlerinin alternatif ölçme değerlendirme tekniklerini kullanmada karşılaştıkları problemler: Trabzon örneği. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi, VI*(1), 122-141.

Creswell, J. W. (2017). *Karma yöntem araştırmalarına giriş* (M. Sözbilir, Trans., Ed.). Ankara: Pegem Akademi.

Darling, Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective Teacher Professionel Development.* Palo Alto, CA: Learning Policy Institute. https://learningpolicyinstitute.org/product/teacher-prof-dev.

Deglau, D., & O'Sullivan, M. (2006). Chapter 3: The effects of a long-term professional development program on the beliefs and practices of experienced teachers. *Journal of Teaching in Physical Education, 25*(4), 379-396.

Demirel, Ö. (2009). Kuramdan uygulamaya eğitimde program geliştirme. Ankara: Pegem Akademi.

Doppelt, Y., Schunn C. D., Silk, E. M., Mehalik, M. M., Reynolds, B., & Ward, E. (2009). Evaluating the impact of facilitated learning community approach to professional development on teacher practice and student achievement. *Research in Science and Technological Education, 27*(3), 339-354.

Gelbal, S. & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 33*(5), 135-145.

Germer, C. K., Siegel, R. D. & Fulton, P. R. (2005). *Mindfulness and psychotherapy.* New York: Guilford.

Giraldo, F. (2014). Bir mesleki gelişim programının İngilizce öğretmenlerinin sınıf performansı üzerindeki etkisi. *PROFİL Öğretmenlerin Mesleki Gelişiminde Sorunlar, 16*(1), 63-76.

Gosser, D. K., & Roth, V. (1998). The workshop chemistry project: Peer-led team-learning. *Journal of Chemistry Education, 75*(2), 185.

Greenleaf, C. L., Hanson, T. L., Rosen, R., Boscardin, D. K., Herman, J., Schneider, S. A., Madden, S., & Jones, B. (2011). Integrating literacy and science in biology: Teaching and learning impacts of reading apprenticeship professional development. *American Educational Research Journal, 48*(3), 647-717.

Gül, E. (2011). *İlköğretim öğretmen adaylarının ölçme-değerlendirme okuryazarlığı ve ölçme-değerlendirmeye ilişkin tutumlarının belirlenmesi* (Unpublished master thesis). Fırat Üniversitesi Fen Bilimleri Enstitüsü, Elazığ.

Gürsoy, G., & Aydoğdu, M. (2020). Effects of contentassociated applied measurement-evaluation course. *Yükseköğretim Dergisi, 10*(1), 96-111.

Heller, J. I., Daehler, K. R., Wong, N., Shinohara, M., & Miratrix, L. W. (2012). Differential effects of three professional development models on teacher knowledge and student achievement in elementary science. *Journal of Research in Science Teaching, 49*(3), 333-362.

Jeanpierre, B., Oberhauser, K., & Freeman, C. (2005). Characteristic of professional development that effect change in secondary science teachers' classroom practices. *Journal of Research in Science Teaching, 42*(6), 668-690.

Kabat-Zinn, J. (2005). *Full catastrophe living: Using the wisdom of your bodyand mind to face stress, pain, and illness (15th anniversary ed).* NewYork, NY: Bantam Dell.

Kağıtçıbaşı, Ç. (2006). *Yeni İnsan ve İnsanlar.* İstanbul: Evrim Yayınevi.

Kılınç, M. (2011). Öğretmen adaylarının eğitimde ölçme ve değerlendirmeye yönelik özyeterlik algı ölçeği. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 12*(4), 81-93.

Kleickmann, T., Trobst, S., Jonen, A., Vehmeyer, J., & Moller, K. (2016). The effects of expert scaffolding in elementary science professional development on teachers' beliefs and motivations, instructional, practices and student achievement. *Journal of Educational Psychology, 108*(1), 21-42.

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

Kruse, L., Impellizeri, W., Witherel, C. E., & Sondergeld, T. A. (2020). Evaluating the Impact of an Assessment Course on Preservice Teachers' Classroom Assessment Literacy and Self-Efficacy. *Mid-Western Educational Researcher*, *32*(2), 107.

Landry, S. H., Anthony, J. L., Swank, P. R., & Monseque-Bailey, P. (2009). Effectiveness of comprehensive professional development for teachers of at-risk preschoolers. *Journal of Educational Psychology, 101*(2), 448-465.

Lee, J. E., & Son, J.-W. (2015). Two teacher educators' approaches to developing preservice elementary teachers' mathematics assessment literacy: Intentions, outcomes, and new learning. *Teaching and Learning Inquiry: The ISSOTL Journal, 3*(1), 47-62. doi: 10.2979/teachlearninqu.3.1.47

Lumpe, A. T. (2007). Research-based professional development: Teachers engaged in professional learning communities. *Journal of Science Teacher Education, 18*(1), 125-128.

McCawley, F.P. (2009). *Methods For Conducting An Educational Needs Assessment. University of Idaho (August).* Retrieved from http://www.extension.uidaho.edu/publishing/pdf/BUL/BUL0870.pdf

McMeeking, L B. S., Orsi, R., & Cobb, B. (2012). Effects of a teacher professional development program on the mathematics achievement of middle school students. J*ournal for Research in Mathematics Education, 43*(2), 159-181.

Meissel, K., Parr, J. M., & Timperley, H. S. (2016). Can professional development of teachers reduce disparity in student achievement? *Teaching and Teacher Education, 58*, 163-173.

Merriam, S. B. (1998) *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass Publishers.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12*(2), 101-113. doi:10.1177/1365480209105575

Mertler, C. A., & Campbell, C. S. (2005, April). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the assessment literacy inventory*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, QC, Canada.

Metin, M., & Özmen, H. (2010). Fen ve teknoloji öğretmenlerinin performans değerlendirmeye yönelik hizmet içi eğitim ihtiyaçlarının belirlenmesi. *Kastamonu Eğitim Dergisi*, *18*(3), 819-838.

Miles, M, B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook.* (2nd ed). Thousand Oaks, CA: Sage.

Mutluer, C. T. (2015). Bolu ilindeki öğretmenlerin ölçme ve değerlendirme genel yeterlilik algılarının incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi, 15*(1), 159-173.

National Staff Development Council (2009*). Ensure great teaching for every student.* NSDC Policy Points, 1 (2).

O'Sullivan, M. C. (2000). Needs assessment for INSET for unqualified primary teachers in Namibia: An effective model. *Journal of Comparative Education, 30*(2), 211-234.

Odo, D. M. (2016). An investigation of the development of pre-service teacher assessment literacy through individualized tutoring and peer debriefing. *Journal of Inquiry and Action in Education, 7*(2), 31-61. Retrieved from https://eric.ed.gov/?id=EJ1133529

Oğuz, E. (2009). İlköğretim okulu yöneticilerinin karar verme stilleri. *Kastamonu Üniversitesi Kastamonu Eğitim Fakültesi, 17*(2), 415-426.

Önen, F., Mertoğlu, H., Saka, M., & Gürdal, A. (2009). Hizmet içi eğitimin öğretmenlerin öğretim yöntem ve tekniklerine ilişkin bilgilerine etkisi: Öpyep örneği. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi, 10*(3), 9-23.

Özbaşı, D., & Çıkrıkçı Demirtaşlı, N. (2013). Sınıf öğretmenlerinin ölçme ve değerlendirme ile ilgili yeterliklere ilişkin algılarının çeşitli değişkenler açısından incelenmesi. Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, *46*(2), 25-46.

Özdemir, S., & Yalın, İ. (2007). *Öğretmenlerin hizmet-içi eğitim ihtiyacının belirlenmesi üzerine bir araştırma.* Paper presented at the 7th National Educational Sciences Congress, 5-7 Eylül, 60-63, Ankara Üniversitesi, Ankara.

Özenç, M. &, Çakır, M. (2015). *Sınıf öğretmenlerinin alternatif ölçme ve değerlendirme yeterliklerinin belirlenmesi. İlköğretim Online, 14*(3), 914-933.

Özsoy, S., & Özsoy, G. (2013). Eğitim araştırmalarında etki büyüklüğü raporlanması. *İlköğretim Online, 12*(2), 334-346.

Patel, J. (2007). *A study of in-service training programmes conducted by DIETs of Guajarat State* (Unpublished doctorate thesis). The Mahara Sayajirao University of Baroda, Vadodara, India.

Pepeler, E., Özbek, R., Adanır, Y., & Kılavuz, Y. (2017). Ortaokul öğretmenlerinin MEB hizmetiçi eğitim faaliyetlerine ilişkin görüşleri (Muş ili örneği). *Journal of Research in Education and Teaching, 6*(1), 71-83.

Quilter, S. M., & Chester, C. (1998, October). *Inservice teachers' perception of educational assessment.* Paper presented at the Annual Meeting of the Mid-Western Educational Researh Assosiation, Chicago.

Quilter, S. M., & Gallini, J. K. (2000). Teachers' assessment literacy and attitudes. *The Teacher Educator, 36*(2), 115-131.

Professional Development Workshops: Determination and Fulfillment of Educational Requirements in Measurement-Evaluation

27

Randel, B., Apthorp, H., Beesley, A. D., Clark, T. F., & Wang, X. (2016). Impacts of PD in classroom assessment on teacher and student outcomes. *The Journal of Educational Research, 109*(5), 491–502.

Reeves, T. D. & Honig, S. L. (2015). A classroom data literacy intervention for pre-service teachers. *Teaching and Teacher Education, 50*, 90-101. doi:10.1016/j.tate.2015.05.007

Resnick, L. B. (2005). Teaching teachers: professional development to improve student achievement. *AERA Research Points, 3*(1), 1-4.

Russell, M. G., Hey, R. N., Thoen, G. A., & Walz, T. (1978). The choice of childlessness: A workshop model. *Family Coordinator, 27*(2), 179-183.

Şanlı, C. & Pınar, A. (2017). Tamamlayıcı ölçme ve değerlendirme gelişim programının coğrafya öğretmen adaylarının yeterlik algısı üzerine etkisi. *Marmara Coğrafya Dergisi, 35*, 21-39.

Sathyanesan, P. (2001). *Effectiveness of in service training programs for teachers and Headmasters by DIETS* (Unpublished doctorate thesis). Mahatma Gandhi University, India.

Shaha, S. H., & Ellsworth, H. (2013). Predictors of success for professional development: Linking student achievement to school and educator successes through on-demand, online professional learning. *Journal of Instructional Psychology, 40*(1), 19-26.

Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers'assessment literacy. *J Sci Teacher Educ, 22*, 371–391.

Smith, M., & O'Day, J. (1990). Systemic school reform. *Journal of Education Policy, 5*(5), 33-267. doi:10.1080/02680939008549074

Stegall, C. L. (2015). *Evaluating the effects of a* professional development *Workshop on the teaching practices of elementary teachers: An action research study* (Unpublished doctorate thesis). Capella University, Minneapolis.

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Sage Publications, Inc.

Süleymangil, M. (2013). *Sınıf içi öğretim etkinliği geliştirme ve kullanma çalışmalarının sınıf öğretmenlerinin mesleki gelişimleri üzerindeki rolü* (Unpublished master thesis). Adnan Menderes Üniversitesi, Aydın.

Supovitz, J.A. & Turner, H.M. (2000). The effects of professional development on science teaching practices and classroom culture. *Journal of Research in Science Teaching, 37*(9), 963-980.

Tabarlet, J. E. (1994). *Teacher implementation of alternative assessment procedures for student learning in selected Texas school districts: An examination of the causal variables* (Unpublished doctorate thesis). Texas A&M University, Texas.

Taymaz, A. H. (1997). *Hizmetiçi eğitim kavramlar ilkeler yöntemler.* Ankara: TAKAV Tapu ve Kadastro Vakfı Matbaası.

Teddlie, C., & Tashakkori, A. (2015). *Karma yöntem araştırmalarının temelleri.* Y. Dede ve S. B. Demir (Çev. Ed.). Ankara: Anı.

Töre, E. (2017). Öğretmenlerin mesleki gelişimi için atölye çalışması modeline dayalı bir uygulamanın geliştirilmesi ve değerlendirilmesi. *Marmara Üniversitesi Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi, 46,* 133-150 Doi: 10.15285/maruaebd.2740

Vidovich, L. (2003). Methodological framings for a policy trajectory study. In Tom O'Donoghue & Keith Punch (Eds.), *Qualitative educational research in action: Doing and reflecting.* London: Routledge Falmer.

Volante L. & Beckett D. (2011). Formative assessment and the contemporary classroom: Synergies and tensions between research and practice. *Canadian Journal of Education 2,* 239-255. Retrieved from https://eric.ed.gov/?id=EJ936752

Watt, H. M. (2005). Attitudes to the use of alternative assessment methods in mathematics: A study with secondary mathematics teachers in Sydney. *Educational Studies in Mathematics, 58,* 21–44.

Yalın, H. İ. (2001). Hizmet içi eğitim programlarının değerlendirilmesi. *Milli Eğitim Dergisi, 150*, 58-68.

Yıldırım, A., & Şimşek, H. (2011). *Qualitative research methods in social sciences.* Ankara: Seçkin Yayıncılık.