

Investigating the Needs for Measurement and Evaluation Course: A Case Study on English Language Teaching Program

Memet Karakuş¹, Buket Turhan Türkan¹

¹Çukurova University, Faculty of Education, Department of Educational Sciences, Adana, Turkey

Correspondence: Memet Karakuş, Çukurova University, Faculty of Education, Department of Educational Sciences, Adana, Turkey.

Received: February 8, 2017

Accepted: March 14, 2017

Online Published: March 21, 2017

doi:10.11114/jets.v5i4.2205

URL: <https://doi.org/10.11114/jets.v5i4.2205>

Abstract

This study aims at investigating the needs related to the attainments, content, processes of teaching-learning and measurement-evaluation aspect of Measurement and Evaluation course. This research was designed as a case study which is one of the qualitative research designs. The most common problem in Measurement and Evaluation courses was found to be the statistics and the source of this problem was students' insufficient mathematical knowledge and skills. It was revealed that there was a need to include more practice, examples and activities in the course. In addition, it was understood that measurement instruments which are peculiar to each discipline need to be included in the scope of the course. Besides, as well as theoretical bases of the course, concrete practices and examples encompassing knowledge and skills that can be used in teaching profession might be included in the course.

Keywords: measurement and evaluation; needs-analysis; higher education; English language teaching program

1. Introduction

Education underpins the success of a society in the future; therefore, it is essential that curricula should be scientific and up-to-date and these curricula should be improved in the light of research and development. Most importantly, in today's world where the expansion of information and the development of technology are so rapid, education has become more crucial. From primary education to university level, the prior aim may be said to help individuals gain necessary knowledge, skills and values.

As in every level of education, it is important to conduct research on how education could be more effective and fruitful in higher education. It is also vital for graduates of higher education institutions to be qualified and intellectual in their study area because one of the significant purposes of higher education is to educate individuals as well-informed and skillful people. Aims, subjects that will be covered and teaching methods should be well-designed in practices involved in higher education which are well-planned in a systematic way. Another point which is as important as this is about measuring and evaluating in a scientific and proper manner. The reason behind this is that the processes of teaching and learning should be reviewed and re-organized with the help of measurement and evaluation practices on a regular basis in order to make the teaching and learning processes more fruitful. Accordingly, the ones who carry out measurement and evaluation practices and who teach how to measure and evaluate have important responsibilities.

Measurement and evaluation is attracting an increasing attention, as teachers have primary responsibility for evaluating teaching and learning (Zhang & Burry-Stock, 2003). According to Haladyna, Downing and Rodriguez (2002), measurement and evaluation should be seemed as an important component of teaching and learning process. Particularly, formative evaluation has an important effect on student development (Rodriguez, 2004). Evaluation has an important role on planning, practicing and developing the teaching process (Bachor & Baer, 2001). For this reason, teaching requires a continuous decision making and evaluation process (Hoover & Abrams, 2013). The current emphasis on evaluation is to draw attention to the goals and strategies of classroom assessment of student learning (Cutsforth, 2013). Measurement and evaluation should include the decisions of the teachers about both formative and summative evaluation (McMillan, 2003). Also, educational measurement and evaluation should not be limited only by the application of measurement principles; it should be examined with the integration of both teaching and classroom management (Brookhart, 1999). Teachers need to have knowledge about how to organize, manage, score and use the results for quality performance assessment (Arter, 1999). Evaluation is a feedback mechanism to improve learning, and

so, learning can be improved by developing teachers' evaluation practices and skills (Zhang & Burry-Stock, 2003). Since limited knowledge about measurement and evaluation will affect the teaching process negatively, it is important to make a special effort about teachers' development (Rodriguez, 2004). In this respect, all the stakeholders in educational area should take responsibility for improving the measurement and evaluation literacy (Mertler, 2005).

There has been a debate over what Measurement and Evaluation course should include from past until today in educational arena. In the middle of 20th century, there was an effort to teach pre-service teachers how to use standard tests and how to prepare their own tests according to the principles of scientific measurement. In that period, course books involved validity, reliability, developing tests, various item types, item-analysis and statistical analyses of test results which are similar to the basis of many course books used today. Later in the 20th century, a perception of quantitative measurement and evaluation focusing on the product has started to be dominant. This perception has transformed into an approach which brings the things understood and achieved by students into forefront with the influence of cognitive and constructivist teaching and learning paradigms and also this perception includes the idea that measurement and evaluation is a formative and integrated part of the learning process (Aktan & Çepni, 2010).

Some studies (Çakan, 2004; Özsevgeç, Çepni & Demircioğlu, 2004; Birgin, 2006; Erdal, 2007, Gelbal & Kelecioğlu, 2007) conducted in Turkey indicated that teachers did not have sufficient knowledge and skills about measurement and evaluation. On the other hand, according to Kilmen and Demirtaşlı (2009), the main reason why teachers cannot carry out measurement and evaluation practices properly is that courses and practices in undergraduate study at university level were not adequate in teaching core knowledge and skills in relation to measurement and evaluation. According to Brookhart (1999), the content of the educational measurement and evaluation course should include different topics than the content of introduction to psychometry courses and the content of the educational measurement and evaluation should include the teaching of concepts, application examples for classroom practices, case studies, simulation and discussion. Since educational measurement and evaluation is integrated with instruction, pre-service teachers should be taught that evaluation is a part of teaching, by focusing on appropriate goals for improving student performance (McMillan, 2003). Teacher educators should draw attention to how to evaluate content-specific goals and analytical thinking (Cutsforth, 2013). It should be ensured that preservice teachers have responsibility for their future students, since teaching the concepts and techniques about educational measurement and evaluation is very important (Mertler, 2005). In the study conducted by Campbell and Evans (2000), it was determined that the pre-service teachers did not follow the recommended evaluation practices in teaching process; it was revealed that more than constructing the knowledge was needed to make the recommended evaluation practices actual and it is stated that there is a need for studies about the elimination of contradictions between teaching and practices about measurement and evaluation for pre-service teachers. Therefore, it is crucial to define the existing situation of Measurement and Evaluation course and its teaching to pre-service teachers and this can be helpful for education faculties to realize the drawbacks if there is any. Moreover, it should be an examination on what the curriculum of a Measurement and Evaluation course which can provide pre-service teachers to gain main concepts and principles of measurement and evaluation and also to gain necessary knowledge and skills while they teach should cover. In line with these explanations and justifications, research questions below are aimed to be answered in this study:

1. What is the existing situation of teaching of Measurement and Evaluation course which is an obligatory course in education faculties?
2. What are the views of students and lecturers about teaching of Measurement and Evaluation course?
3. What are the needs in relation to teaching of Measurement and Evaluation course?

2. Method

This study was designed as a case study which is one of the qualitative research designs. It is important for case studies to investigate one case or more cases with an appropriate instrument deeply, extensively and in detail (Punch, 2005). Case studies are restricted to some certain features have particular and characteristic qualities (Freebody, 2004). In this study, what is investigated was a case of Measurement and Evaluation course in an English Language Teaching Department at a university located in the southern part of Turkey. The main reason why this department was selected as a case is that Measurement and Evaluation course contained mathematical and statistical calculations and pre-service teachers in this department were not so much exposed to mathematical and statistical calculations in high school and undergraduate study at the university. For this reason, it was presumed that the existing situation and problems of Measurement and Evaluation course in this study area can be revealed more effectively. In addition, results linked to similar attainments, content and teaching methods of Measurement and Evaluation course in all departments can be found out.

2.1 Participants

To identify participants of the study, criterion sampling method which is one of the purposive sampling methods was used. Purposive sampling is a method designed to develop an approach or theory or to deeply understand experiences of selected individuals or groups (Devers & Frankel, 2000). The researcher has an active role in selecting the most appropriate sampling group in order to give answers to research questions (Marshall, 1996) and he/she selects individuals or groups who can be rich sources of information related to the research question (Devers & Frankel, 2000). In that respect, the criteria for defining the pre-service teacher participants are being a volunteer, having taken the Measurement and Evaluation course, completing the second semester of their third year in English Language Teaching department. As for defining lecturers, the criteria were being volunteer and having taught Measurement and Evaluation course at least one semester in English Language Teaching department. Participating pre-service teachers were composed of 7 males and 12 females. 14 of them graduated from Anatolian High Schools, 3 of them graduated from Anatolian Teacher Training High School and 2 of them graduated from General High School. Regarding participants' general average grades, one of them had 3.51-4.00, 10 of them had 3.01-3.50, 4 of them had 2.51-3.00 and 4 of them had 2.01-2.50. As to participating pre-service teachers' passing grades, 7 of them had AA, 5 of them had CC, 3 of them had BA, one of them had BB and one of them had CB. In terms of lecturers, 2 of them were females and one of them was male. All of them were working in the Curriculum and Instruction program which is a part of the department of Educational Sciences. 2 of them were assistant professors and one of them was an instructor.

2.2 Data Collection Tools and Data Analysis

To collect data, semi-structured interviews were conducted. Two different semi-structured interview forms were prepared for pre-service teachers and lecturers. Experts' opinions for the interview forms were obtained and in the light of experts' opinions, necessary alterations were made in the interview forms and pilot study was carried out afterwards. As a result of these, interview forms were re-shaped.

In the first part of the interview forms, there are demographic information questions. Following this, in the interview form of pre-service teachers; in line with the attainments, content, teaching-learning processes and measurement-evaluation process of Measurement and Evaluation course, questions that can define the existing situation and some other questions that can reveal suggestions for improving Measurement and Evaluation course are included. In the interview form of lecturers, there are questions similar to the questions in the interview form of pre-service teachers and also there are more questions related to defining a need for a new curriculum and teaching the course in the same way in all departments.

Interviews were conducted between July and September in the academic year 2015-2016. All interviews were carried out with the help of audio-recordings; for this, permissions from participants were provided.

Data obtained from interviews were analyzed with the help of content analysis method. The aim of content analysis method is to reach themes and patterns which can explain data collected to reveal meanings and main contexts in the qualitative data (Patton, 2002). Accordingly, stages suggested by Yıldırım and Şimşek (2008) were taken into consideration and audio-recordings were transcribed on computer in the content analysis process. In this way, raw data were constructed and the data were coded by conceptualizing. Then, themes and sub-themes explaining these codes were created by categorizing the codes according to their characteristics. NVIVO 10 was utilized in the analysis of data. To present findings, pre-service teachers were assigned different codes such as EPT1, EPT2 and so on. And the lecturers were coded as L1, L2 and the like. In the presentation of quotations, the source of data was specified with these codes.

For the sake of validity and reliability, triangulation and peer examination methods suggested by Merriam (1995) were used. To triangulate the data, there were different data sources and these sources were both pre-service teachers and lecturers. As for peer-examination, the co-researcher was involved in the detailed examination of codes, themes, sub-themes and findings. This co-researcher was also an expert in curriculum and instruction and she observed Measurement and Evaluation course throughout two years. In accordance with that, peer opinion in relation to the appropriateness of the codes and their consistency with themes and sub-themes was received. Moreover, peer opinion was also received in terms of whether the findings reflected the truth and whether the comments were reasonable or not. In this process, inappropriate and inconsistent points were examined and discussed in detail and findings were organized by reaching a common ground. What is more, for the sake of validity in qualitative research, Yıldırım and Şimşek (2008) proposed the researchers to report the data in detail and to explain how they reach their results. In parallel with this notion, to have valid results, the data were described elaborately, direct quotations were illustrated and explanations about how results were obtained were given.

3. Findings

Findings were presented in categories of attainments of Measurement and Evaluation course, course content,

teaching-learning processes and measurement and evaluation aspects respectively. Accordingly, firstly the pre-service teachers' views and secondly the lecturers' views in relation to each category were introduced.

3.1 Findings about Attainments of Measurement and Evaluation Course

Views of pre-service teachers in relation to the attainments of Measurement and Evaluation course were categorized under four themes which were knowledge of aims and attainments of the course, expectations and perceptions before the course, meeting the course's expectations and influence of attainments on professional life. Two sub-themes were generated and they were attainments about measurement (f:29) and attainments about evaluation (f:18). In the sub-theme of attainments about measurement, there were codes such as measurement in the right way (f:11), preparing measurement instruments (f:9), analysis of tests and items (f:5), measurement techniques (f:3) and defining the validity of tests (f:1). As for the sub-theme of attainments about evaluation, there were codes such as evaluating in the right way (f:13), defining criteria (f:3), defining the readiness (f:1) and examining the process of development (f:1). Concerning this theme, EPT12 stated his/her opinions by saying *"Our aim is to learn about test types in this course and also to gain skills which can provide us to learn how to evaluate and analyze the tests. Moreover, another aim is to describe the purposes of the tests and to realize why we apply the tests."* This quotation shows that EPT12 talked about measurement techniques, preparing measurement instruments and evaluating in the right way.

In the scope of the theme of expectations and perceptions before the course, two sub-themes were created as expectations before the course (f:17) and perceptions before the course (f:8). The codes under the sub-theme of expectations before the course can be listed as learning how to conduct measurement and evaluation practices (f:6), learning how to prepare a test (f:6), learning mathematical and statistical calculations (f:3) and learning the concepts of measurement and evaluation (f:2). In terms of the sub-theme of perceptions before the course, there were codes such as accepting the course as an undetailed course (f:3), accepting the course as a verbal course in which there are not so many mathematical calculations (f:3), accepting the course as a difficult course (f:1) and accepting the course as having special measurement techniques about the profession (f:1). Within this theme, EPT14 said that *"I was expecting to learn how to test. I mean I was supposing that I would learn how to prepare tests and how to test students' knowledge. I did not know anything about these mathematical-statistical analyses and I was scared about them because I am not good at mathematics. Apart from these, my expectations were met."* That is to say, EPT14 mentioned about his/her expectations about learning how to test and evaluate along with learning how to prepare tests. He/she also expressed that he/she perceived the course as not including so many mathematical calculations before the course.

Students stated that their expectations were met from the Measurement and Evaluation course in varying ratios between 50% and 100%. For this reason, it can be said that the course met the students' expectations. Furthermore, they asserted that the course was beneficial in terms of learning how to prepare a test (f:2), contributing to KPSS (Public Personnel Selection Examination) (f:1), overlapping with course called "testing" (another course taught in the department) (f:1), learning more than what is necessary for professional teaching life (f:1). EPT4 indicated that the course met the expectations 90% and added *"I was just thinking about whether the content of the course should be defined by students or it should be based on a source. The course included what I expected in a ratio of 90%. The remaining 10% part is about some imaginary things that I thought about."*

Regarding the theme of influence of attainments on professional life, there were three sub-themes as evaluating the teaching process (f:16), preparing and implementing tests (f:13) and negative views (f:18). Under the sub-theme of evaluating the teaching process, some codes such as measurement and evaluating in the right way (f:8), defining the level of learning and success (f:3), defining the effectiveness of teaching (f:1), making planned evaluations (f:2), defining the reasons of unsuccessful aspects (f:1) and learning how to evaluate the process (f:1). The codes under the sub-theme of preparing and implementing tests were as defining validity and reliability (f:6), organizing content validity (f:2), generating questions and classifying them according to attainments (f:3) and identifying and implementing tests (f:2). When the sub-theme of negative views were examined, it is seen that there were two codes which are about not implementing everything learnt (f:9) and being of the opinion that what is learnt is not used in the existing system (f:9). About the contributions of the course to the professional life, EPT15 said that *"I learnt how to evaluate prepared questions that I found from somewhere else....I already know that if each question is difficult or easy in a test, this will decrease the distinctiveness of the test. I am sure I will get help from this knowledge when I prepare a test and make evaluation."* Contrary to these positive views, EPT9 was in the idea that what is learnt in the course is not used in the existing system by saying that *"...but when I think about the past, we see that teachers did not use any of these..."*

Lecturers' views in relation to the attainments of the course were categorized under the themes of knowledge and skills aimed to be gained (f:13), ways for defining overall aims and attainments (f:5) and influence of attainments on professional life (f:6). In the sub-theme of knowledge and skills aimed to be gained, lecturers emphasized such skills as preparing, implementing and analyzing measurement instruments (f:3), measurement in the right way (f:2), knowing the

main concepts (f:2) and measurement techniques (f:1), evaluating in the right way (f:3), controlling whether the attainments can be achieved (f:1) and realizing that the elements of a curriculum is a whole (f:1). With regard to this, L1 underlined evaluating in the right way, preparing and implementing measurement instruments by saying that *“Examining exam papers, evaluating the scores, transforming the raw scores into grades, ranking students according to their success are the things that they will use in their future career. They will use some techniques while preparing questions, for instance; not using the options of none of them, at every time, generally in certain questions...”*

In the scope of ways for defining overall aims and attainments, lecturers stated that they organized the content of their courses by getting help from different areas (f:3) and taking needs into consideration (f:2). In parallel with this, lecturers used the main aims determined by Council of Higher Education (f:2), the constructs in the course books (f:1), approaches in the available curriculum (f:1) and making connection with KPSS (f:1) as the base. As to taking needs into consideration, they referred to knowledge and skills that will be used in teaching profession (f:1) and the level of students (f:1). In the light of this theme, L3 explained the sources he/she got help from by saying that *“I think that it is necessary to maintain this course and its general scope in education faculties across our country...I try to give importance to main components defined in teacher education programs in education faculties. Plus, I make an empathy to understand what kind of curriculum is implemented at the moment, what kind of a process is included in this curriculum, what sort of teaching-learning approach it contains and what it suggests teachers to do in terms of measurement and evaluation.”*

Pertinent to influence of attainments on professional life, lecturers stressed the aspects of evaluating the process of teaching (f:3), preparing measurement instruments (f:2) and defining and organizing content validity (f:1). In the interview, L2 emphasized the importance of learning how to prepare measurement instruments by saying that *“Students learn to prepare tests... for example, while I was teaching the ways for increasing the validity and reliability..., they said that most of the teachers in the university did not apply the rules that I taught. I said that they might not apply but this does not necessarily mean that you will not apply.....in the future, your students may say that our teacher does not apply the necessary examination rules in primary school....in high school....so, I say that you need to apply the rules in order to prevent negative feedback from your future students.”*

3.2 Findings about the Content of Measurement and Evaluation Course

Views of pre-service teachers on the content of Measurement and Evaluation course were gathered under seven themes which were efficiency and necessity of the content, important subjects, difficult subjects and the reasons, subjects that can be added, subjects that can be omitted, suggestions for the content and adequacy of the time for the course. The theme of efficiency and necessity of the content included two sub-themes as necessity of the content (f:19) and efficiency of the content (f:15). In the scope of necessity of the content, there were codes such as necessity of all units and subjects (f:13), necessity of only main subjects (f:3) and the unnecessary of advanced statistical calculations (f:3). In terms of efficiency of the content sub-theme, 15 pre-service teachers uttered that they accepted all units and subjects as efficient as important and 4 of them did not state any opinion about this issue. EPT12 stated that he/she accepted all subjects as necessary and said *“ I think all of them are necessary....If we apply a test without knowing whether it is valid or reliable, we cannot evaluate students' real skills....All subjects should be included in this course because I could not omit any of them from the content and we use all test types in English courses....Apart from national exams conducted by ÖSYM (Student Selection and Placement Center), we have to know about all test types; for this reason, we should learn all of them.”*

In the theme of important subjects, there were the sub-themes of statistical subjects (f:5), main subjects (f:4), measurement instruments (f:4) and qualities of test instruments (f:2). Regarding the sub-theme of statistical subjects, normal distribution curve (f:1), difficulty of the test (f:1), item distinctiveness (f:1), distractors, reliability, mean score (f:1) and focusing more on statistical analysis (f:1). As for main subjects, there were codes of subjects about evaluating and interpretation (f:3) and main concepts about measurement (f:1). When the sub-theme of measurement instruments are examined, it is seen that there was the code of test types (f:4) and in the sub-theme of qualities of measurement instruments, codes of interpreting validity and reliability without mathematical calculations (f:1) and content validity (f:1) were seen. About this theme, EPT15 gave importance to test types by saying *“Question types in the measurement and evaluation, written or oral questions, should be taught, at least to introduce students into teaching profession.”*

For the theme of difficult subjects and the reasons, pre-service teachers pointed out that they had difficulties in calculating the KR-20 value because of not putting the formula into practice (f:3), not studying the course regularly (f:1) and insufficient mathematical knowledge and skills (f:15). They had difficulty in putting the formula into practice because of item-difficulty index (f:1) and standard deviation (f:1). Also, they experienced some obstacles in coefficient of variance (f:1) because of not studying the course regularly and there were some other obstacles in interpreting tables and graphics (f:1) because they involve numerical information. Relatedly, EPT13 expressed his/her opinions by saying

“The most difficult part of the course was calculating fractions and square root but I realized that we can do them....I just answered two mathematic questions in university entrance examination because my background knowledge in mathematics is limited...”

Under the theme of subjects that can be added, two sub-themes were defined as subjects related to the subject area (f:10) and general subjects (f:4). Considering the sub-theme of subjects related to the subject area, pre-service teachers thought that subjects about measuring speaking, writing, listening skills (f:8) and examining/developing tests related to the subject area (f:2) could be added to the content. In addition, for the sub-theme of general subjects, they stated that they could be evaluated with more projects and performance tests in the course (f:2) and some subjects related to KPSS (f:2) could be added. As for subjects that can be omitted, pre-service teachers argued that statistical calculations (f:6) and calculations of KR-20, KR-21 values (f:3) could be omitted. When the theme of adequacy of the time for the course is considered, 11 students thought that the allocated time for the course was sufficient and 6 of them believed that the time for the course was not adequate. And 2 students did not take any position about the issue. Regarding this theme, EPT10 addressed to some aspects about measuring writing, speaking, listening by saying *“It should not only be oriented to testing, as English language teachers we can only test grammar knowledge but we have writing, listening and speaking sections and these sections may not be tested. Maybe at the moment, this course could be taught to a Turkish or Mathematics teacher in this way.....but it would be logical to learn how to measure writing, listening and speaking skills.”*

In relation to suggestions for the content theme, three sub-themes were put together and they were subjects that should be given importance (f:7), suggestions for the intense content (f:2) and other suggestions (f:2). As for the sub-theme of subjects that should be given importance, such codes as main subjects that will be used in teaching profession (f:2), verbal information and subjects (f:1), classical written exams (f:1), mathematical subjects (f:1), subjects for generating one's own measurement techniques (f:1) and subjects for how to use what is learnt (f:1) were involved. In the scope of suggestions for the intense content, there were suggestions of teaching the course throughout two semesters (f:1) and increasing the time for the course (f:1). Other suggestions sub-theme contained the codes of organizing the content based on different departments (f:1) and reviewing the Turkish equivalents of terms (f:1). Concerning this theme, EPT19 stated *“...This course is available in every department but our department may be ignored and it is upsetting that Measurement and Evaluation course taught to elementary or computer teachers are the same with English language teachers....”* That is EPT19 suggested that the content should be altered depending on the departments.

Views of lecturers on the content of the course were categorized into seven themes of ways for defining the appropriate content, subjects difficult for students, subjects that can be omitted, subjects that should be included in the content, adequacy of the time for the content and the reasons, problems about the content, suggestions for the content. In the scope of ways for defining the appropriate content, there were the sub-themes of getting help from the existing content (f:5), based on inquiry (f:3) and approaches based on organizing the main content (f:3). In the sub-theme of getting help from the existing content, codes of content determined by Council of Higher Education (f:2), course books related to the subject area (f:2) and having a well-developed content (f:1) were available. Within the sub-theme of based on inquiry, there were codes which were paying attention to current developments (f:1), paying attention to students' characteristics (f:1) and organizing the content according to feedback from students (f:1). In regard to this theme, L1 highlighted how important it was to take the needs into consideration by saying *“...We do not define the content ourselves, we just simplify the content which is already existing according to the departments of our students,....we add or omit subjects with the help of different sources and we pay attention to characteristics of students and the departments.”*

Within the scope of difficult subjects for students, there were two sub-themes which were mathematical subjects (f:5) and main concepts (f:2). The sub-theme of mathematical subjects involved the codes of mathematical and statistical subjects (f:1) and absolute zero and relative zero (f:1). And in the sub-theme of main concepts, there were the codes of intermingling similar concepts (f:1) and intermingling predictive validity and convergent validity (f:1). Based on this theme, L2 mentioned about the difficulty that students had in mathematical calculations by saying *“The most difficult thing for them is subjects that they need to do calculations but this is not because of the difficulty of that subject....Instead it is because of their insufficient mathematical knowledge and skills. For example, students may forget what square root means...”*

Regarding the subjects that can be omitted, there were two codes as statistical subjects (f:2) and the current subjects (f:1). About statistical subjects, there were the codes of formulas and calculations that can be done by software (f:1) and KR20, KR21 values (f:1). Associated with this theme, L1 thought that subjects that are not used today should be omitted by saying *“...There are some question types that are no longer usedwe try to teach students what is important while making an oral examination in 1970s. Ministry of National Education do not support this, instead it suggests us to evaluate students according to their performance in the course...If we have a different educational approach from the one used in 1970s, then we need to omit these.”*

In the theme of subjects that should be included in the content, the codes were qualities of measurement instruments (f:3) and statistical subjects (f:2). In association with qualities of measurement instruments, there were codes of qualities of a good measurement instrument (f:1), consistency (f:1) and distinctiveness (f:1). As for statistical subjects sub-theme, there were main statistical subjects (f:1) and focusing more on the interpretation rather than statistical calculations (f:1). Pertaining to this theme, L3 talked about the need to concentrate more on interpretation about statistical subjects by saying *“We can use package software instead of making calculations on the board...I think that we need to focus more on interpretation by giving examples from real-classroom environments. We can just teach the logic behind calculating mode-median-mean and we can talk about calculations that are already ready on the board. In this way, we will need less time to solve problems in the class.”*

In terms of adequacy of the time for the content and the reasons, there were two sub-themes as finding the time adequate (f:2) and finding the time inadequate (f:1). Under the sub-theme of finding the time adequate, there was the notion that the adequacy of time depended on the teaching method and the inadequacy of time could be related to the intense content and resistance of students. Referring to this theme, L3 pointed out that the time was not adequate because of the intense content by saying *“...if I think about cognitive and emotional knowledge and skills, I believe that the content is too intense, particularly the content about knowledge...the time is for each course 45 minutes and we have three hours a week but it is not adequate unfortunately.”*

The sub-theme of problems about the content contained the codes of students' inability to make sense of the subjects (f:1) and intermingling foreign terms (f:1). Relatedly, L1 touched on students' inability to make sense of the subjects by saying *“They have problems in making sense of the subjects”*.

In the theme of suggestions for the content, two sub-themes were constituted as suggestions for organization (f:5) and suggestions for subjects (f:2). In terms of suggestions for organization, such suggestions as organizing the content based on the departments' structure (f:1), cooperation with departments to organize the content (f:1), getting students' views on organizing content (f:1), organizing content with the help of spiral approach (f:1) and suggestions about problem-based or project-based approaches (f:1). As to suggestions for subjects, repetition of main subjects about curriculum development (f:1) and omitting unnecessary subjects from content (f:1) were proposed. In relation to that, L1 suggested that *“Firstly, some unnecessary subjects should be eliminated and secondly, repetition of main subjects about curriculum development and thirdly, organizing content according to the departments with the other lecturers should be taken into account...”*

3.3 Findings about Teaching-learning Processes of Measurement and Evaluation Course

Views on teaching process of Measurement and Evaluation course were categorized into five themes. In the scope of the first theme which was the existing situation of teaching-learning processes, there were four themes which were the approaches, methods and techniques focused (f:53), variables of teaching process (f:12), concerns for teaching (f:6) and helping students be central and improve (f:4). Within the sub-theme of approaches, methods and techniques focused, the codes of giving examples and making exercises (f:15), presentation by teachers (f:13), lecture (f:12), question and answer (f:7) and discussion (f:4) were presented. As for the sub-theme of variables of teaching process, such codes as giving feedback (f:4), attracting attention (f:2), repetition (f:2), motivation (f:2) and informing about the aims (f:2) were found. In the sub-theme of concern for teaching, there were two codes and they were making effort until the content is totally comprehended (f:5) and overemphasizing the content (f:1). When it comes to the sub-theme of helping students be central and improve, the codes were making students involved in the course actively (f:1), giving importance to students' opinions (f:1), making students gain a critical point of view (f:1) and taking students' competency into consideration (f:1). In relation to this theme, EPT14 mentioned about the techniques that lecturers use and lecturers' concerns of teaching by saying *“Our teacher used PowerPoint presentation, gave so many examples and tried to teach the subjects that we did not understand again and again. I like him/her teaching in this way. Presentation usage was a good idea because it would be time-consuming to write statistical things on the board. We directly focused on our subject without wasting time so I think PowerPoint slides were beneficial. Also, I think that our teacher taught the content in a very good way.”*

When the theme of communication atmosphere of the course is considered, it is seen that there were the codes of positive communication between teacher and student (f:2), student talk with each other (f:1), distraction of the course with the questions (f:1), disrupting the course because of antipathy for mathematics (f:1), lecturers' unchanging tone of voice and lack of body language use (f:1).

For the theme of problems about teaching, there were two sub-themes which were problems related to students (f:5) and problems related to the course (f:3). The codes of the sub-theme of problems related to students were passive behaviors of students (f:3), student talk with each other (f:1) and unwillingness to make exercises about statistical things (f:1). As to the sub-theme of problems related to the course, the codes were insufficient break times (f:1), being unable to teach

some subjects because of time limit (f:1) and inadequacies in prerequisite courses about education (f:1). About this theme, EPT19 uttered *“There was some problems in participation in the course, we could not interpret what was taught and the subjects were already confusing for us...we were so confused with negatively skewed and positively skewed graphics but we tried to do our best to respond to our teacher....”*

In the theme of suggestions for teaching process, there were five sub-themes. The first one was suggestions for making practice and there were the codes of making more practice (f:8), developing tests together (f:6) and giving more examples in the course (f:4). The second sub-theme was suggestions for methods and techniques involved the codes of helping students be active (f:3), making subjects concrete (f:2), making students remember prerequisite knowledge with the help of question-answer and brainstorming (f:2), examining applied tests in the course (f:2), group works (f:2), asking students to present examples (f:1) and integrating enjoyable activities which can decrease the concern of mathematics (f:1). The third sub-theme was suggestions for technology and materials and it contained the codes of using smart board (f:2), getting help from technology for mathematical calculations (f:1), using statistical software (f:1), using the board effectively instead of using slides (f:1) and making obligatory to use course books (f:1). Within the fourth sub-theme which was suggestions for making exercises, there were some codes such as solving more problems (f:5), solving questions involved in the exam for being a teacher (f:2) and providing students to solve problems on the board (f:2). The fifth sub-theme was suggestions for affective domain and it was composed of the codes which are providing all students to participate in the course (f:1), attracting the attention and being motivated (f:1), building positive communication between teacher and student (f:1) and removing students' prejudices and making them motivated (f:1). EPT13 suggested that there should be more practice in teaching-learning processes and said *“There should be more practice because I realized that I could understand better as long as I made practice rather than just looking at what was taught on the board. There was one point that I could not understand and I made practice about it. In this way, I understood it in a short time and I was really shocked about it.”*

Views of lecturers on teaching-learning processes were categorized under five themes which were the existing situation of teaching-learning processes, features of the environment, materials used, problems about teaching and suggestions for teaching. In the sub-theme of the existing situation of teaching-learning processes, there were the codes of approaches, methods and techniques used and activities carried out. In the sub-theme of approaches, methods and techniques used, there were codes which are lecture (f:3), discussion (f:3), question and answer (f:3), projects (f:3), case-study (f:2), cooperative learning (f:2), discovery approach (f:1), expository approach (f:1) and problem-solving (f:1). As for the sub-theme of activities carried out, codes were classifying KPSS questions according to the subjects (f:1) and examining articles (f:1). Relatedly, L3 shared his/her opinion by saying *“In the first weeks, I am active in the subject of main concepts, I mean for the subjects of question and answer, discussion and lecture....The second subject is validity and reliability and I use lecture method to teach them. Firstly, I give prerequisite knowledge and then, I have to give so many examples and I have to use question and answer and discussion techniques....In the third part, if the number of students in the course is appropriate and I can manage the class actively, I start to make students work in small groups... When we start statistical subjects, I use lecture, discussion, reasoning techniques. That is, we focus on reasoning by thinking about how to solve a problem, how to increase mean scores and so on.”*

In the theme of features of the environment, codes were irremovable chairs that are not suitable for group activities (f:2), broken projections (f:1), unsuitable temperature because of broken air conditioner (f:1), dirty places which is sun-drenched (f:1). About this theme, L2 talked about some negative things about the physical environment by stating *“...I never teach a course in a good classroom. I mean the classrooms were either dirty or dusty or sun-drenched and projection was broken or PowerPoint slides were not visible enough.”*

Concerning the theme of materials used, codes were lecture notes (f:1), course books (f:1), worksheets (f:1), articles about the course (f:1), table of specifications (f:1), different scales (f:1), midterm exam scores (f:1) and national exams (f:1) and slides (f:2). In relation to this, L1 said *“We use measurement instruments which are valid and reliable such as exams conducted by ÖSYM....We calculate mode, median of our own exams like midterm and final exams....”*

The sub-theme of problems about teaching contained the codes of unwillingness to take responsibility (f:1), completing assignments easily (f:1), interference with each other (f:1), problems about participation to the lesson (f:1), problems about classroom management (f:1), having difficulty in making practice because of intense content (f:1) and having difficulty in finding examples about the subject area (f:1). In association with this theme, L3 emphasized the unwillingness of students in taking responsibility by saying *“If students are expecting that only the teacher is responsible for the lesson, they do not want to have responsibility. That is, I conflicted with their expectations...”*

In the theme of suggestions for teaching, there were codes of making various exercises and diversifying the teaching methods (f:1), portfolios (f:1), making the lesson enjoyable (f:1), analyzing the questions in the course book (f:1), explaining how to use measurement in teaching profession (f:1), examining applied tests (f:1), analyzing case studies

(f:1), activities to improve skills for preparing questions (f:1), making adjustments with the experts from the field (f:1) and integrating the course in the fourth year of study (f:1). L3 gave some suggestions about observing teachers in the school in terms of measurement and evaluation by saying *“Small projects or research for leading students to real school environments can be used. They can observe teachers in terms of how they prepare questions or what kind of problems they experience. In this way, they can be involved in the process...”*

3.4 Findings about Measurement and Evaluation Process of the Course

Views on this aspect were categorized into four themes. The first theme was the existing situation of measurement and evaluation processes of the course and all students mentioned about being evaluated by a project based on practice rather than a midterm exam. They also said that final exam could be multiple-choice type test and both theoretical and practical knowledge should be evaluated. Other views about this were support from lecturers out of the class (f:2) and providing beneficial attainments for KPSS (f:2). Relatedly, EPT5 pointed out that *“Midterm exam questions were composed of subjects related to practice about subjects that would be covered throughout the semester. We went to a school and applied a test to students. Then, we examined every detail which we learnt in the course and prepared reports about them. This was evaluated as the midterm exam. In the final exam, we responded multiple-choice type questions.”* In the scope of the second sub-theme which was efficiency of measurement and evaluation, there were such codes as finding process-based evaluation more beneficial (f:6), preferring to apply the midterm evaluation as an exam (f:3), finding assignments insufficient (f:3), finding to divide assignments into weeks more effective (f:2), finding practice more effective than exams (f:1), finding product-based evaluation more effective (f:5), being content about multiple-choice type tests in the final exam (f:2), finding product-based evaluation insufficient (f:1) and finding the exam insufficient (f:1). EPT8 thought that process-based evaluation was more effective and said *“Actually, I think that assignments are given in order for us to learn more because everything focused on by the teacher is involved in the assignment. We need to learn the subjects covered to complete the assignment. We reviewed the subjects at home and then we completed the assignment. In this way, we both read and made practice.”* The third sub-theme was students' readiness and it composed of the codes of insufficient readiness for the mathematical calculations (f:14), concern for the mathematics (f:7) and prejudice about the difficultness of the course (f:2). EPT11 spoke of insufficient readiness for the mathematical calculations by saying *“I could understand it if it was more verbal...we always tried to deal with calculations and I could not comprehend the things completely. Because of the fact that we are language students, we are a bit far away from mathematics. Right now, we do not remember how to make subtraction and how to sum up with square roots.”* The last sub-theme was suggestions for measurement and evaluation processes of the course involved the codes of evaluating the performance in the course (f:7), preparing portfolios by dividing the midterm assignment into weeks (f:5), applying tests in the midterm exam and applying projects in the final exam (f:2), making practice in the final exam instead of tests (f:5), involving different test types in the final exam (f:3), applying other test types and evaluating them in the midterm exam (f:1), defining the individual contribution to the midterm assignment (f:1), preparing midterm assignment as individual work rather than group work (f:1), using different test types in the midterm exam (f:1), tests in midterm assignment prepared by students (f:1), using both projects and tests in the midterm and final exams (f:1), making quiz exams (f:1), using supplementary measurement instruments together (f:1) and helping students out of the class (f:1). Regarding this theme, EPT1 suggested that performance in the course should be evaluated by saying *“Maybe attitudes can be taken into consideration for the evaluation.”* and EPT19 suggested that midterm assignment could be divided into weeks by saying *“I think assignments are not appropriate for the nature of this course, assignments should be given as weekly practice of subjects covered...”*

Views of lecturers on measurement and evaluation process of the course were categorized under four themes. The first theme was the existing situation of measurement and evaluation processes and it encompassed the codes of multiple-choice test (f:3), project tasks related to test and item analysis (f:2), assignments (f:2) and classroom practices (f:2). In terms of this theme, L1 asserted that multiple-choice tests were used in measurement and evaluation by saying *“I evaluated students with simple multiple-choice tests; I know that this is not true but I had 475 students in total last year.”* In the scope of second sub-theme which was efficiency of measurement and evaluation process, such reasons were presented as finding multiple-choice type tests easy and practical, finding to use only multiple-choice tests insufficient (f:1), exams that are not reflecting the performance (f:1), finding classroom participation important (f:1), not being able to observe projects made out of class (f:1), not being able to evaluate skills (f:1) and being obliged to use multiple-choice tests because of big number of students (f:1). L2 stressed that exams did not reflect the real performance of students by saying *“I have students who participate in the lesson actively, give right responses or ask questions. I examined exam results of these students and I see that they got low grades....and I was shocked.”* Within the sub-theme of readiness of students, lecturers emphasized insufficient mathematical background of students (f:3), fear of mathematics (f:3) and prejudice and worry about the course (f:2). In addition of these, in relation to readiness of students, codes mentioned were that readiness of students depended on the course of Teaching Principles and Methods

(f:1) and high readiness of students graduated from Anatolian Teacher Training High Schools. Concerning this theme, L1 underlined prejudice and worry about the course along with fear of mathematics by stating that “...*Students have already negative opinions because 80% of departments that I teach have serious negative situations about mathematics. Let's not say they have difficulties most in mathematics but they are fearful about statistical calculations....*” Suggestions for measurement and evaluation process which was the fourth theme were constituted of codes as evaluating skills and practices (f:2), giving immediate feedback to assignments (f:1) and giving importance to classroom participation (f:1). Accordingly, L3 suggested making evaluations for skills and practices by saying “...*I evaluate skills but I wish I could be with students while they make practices and give feedback to them immediately. I wait for the following weeks to give feedback but it would be more useful to give feedback in the classroom after practice. In fact, the things that we do are not wrong; I think that necessary things are not done as it should be. This is the problem.*”

L1 and L3 stated that it was wrong to teach the course in the same way in different departments in the faculty. The reasons for that were about the fact that each teaching profession had its own distinctive measurement techniques. They also said that they had difficulty in giving examples about the subject area in the lesson and for this reason; they needed to cooperate with other departments. In terms of teaching Measurement and Evaluation course in the same way in all departments, L1 said “*Because each teacher will not test the same thing, one of them will test physical performance and another will test cognitive performance....., all these require a different measurement technique. In the first seventh-eighth weeks of this course, main concepts about measurement and evaluation need to be taught in all departments and then measurement instruments appropriate for the departments should be taught.*”

In addition to problems above, lecturers were asked whether they used a syllabus for Measurement and Evaluation course and three of them stated that they used a syllabus. In relation to this, L1 and L3 pointed out that they used the framework of Council of Higher Education as base and they organized this framework according to departments. What is more, L1 said that he/she essentially taught the common subjects, organized the framework according to students' characteristics and got help from course books while preparing the syllabus. L2 indicated that he/she made use of experiences from previous courses to organize the syllabus. In this respect, L3 said “*I follow the framework of Council of Higher Education but the framework changes in time because I make alterations based on the departments. Materials and examples I use may change depending on the departments. That is, when I teach a physical education class, I make use of visuals or checklists that can test psychomotor skills and give suitable examples. When I teach a language class, I collect examples about alternative techniques for measuring vocabulary knowledge from the Internet.*”

Besides, lecturers were asked whether they needed a new curriculum or not. L1 and L2 stated that they did not. The reason behind this was that the course had a structured curriculum according to them. Additionally, L1 said that the curriculum should be revised in the light of current developments rather than developing a new curriculum. L2 said that he/she was not affected by the insufficiency of the curriculum because he/she took courses from experienced teachers. In this regard, L1 stated “...*I can say different things about other courses but Measurement and Evaluation course has a more structured curriculum. Revising the curriculum or re-revising it....portfolios, rubrics should be come into prominence. There are still some books that do not contain these subjects. How to use computers or smart boards for measurement and evaluation should be taught. There should be a curriculum which is renewed and updated according to the needs rather than a completely revised curriculum.*”

4. Conclusions, Discussion and Suggestions

Research findings show that students were aware of aims and attainments of Measurement and Evaluation course and they thought that the course met their expectations. They also stated that the course would provide beneficial contributions to their professional life in a positive way. Moreover, some students were concerned about to what extent teachers were using knowledge and skills which they gained in the course. It can be said that lecturers and students had similar opinions about aims and the attainments of the course. Before the course, it was found out that students had partly negative opinions about the course and the effect of the course on their professional life. In the study conducted by Çobanoğlu and Çepni (2010), lecturers was found to give importance to preparing tests, question types and standards but they thought that item-analysis and calculating reliability were unnecessary subjects. Duban and Küçükylmaz's (2008) study showed that lecturers used different implementations in the process of measurement-evaluation and only a few of them used alternative measurement-evaluation methods in the class. Gullickson (1987) conducted a study on what needed to be taught in Measurement and Evaluation course with teachers and experts and it was revealed that there were significant differences between views on statistical subjects. Again in this study, it was emphasized that the content of the course should be defined in the light of teacher needs. In the study of Gelbal and Kelecioğlu (2007), it was explored that teachers generally used traditional measurement and evaluation methods and they never used other methods such as self-evaluation. Another study conducted by Şahin and Abalı-Öztürk (2014) indicated that pre-service teachers were of the opinion that measurement and evaluation practices in primary schools were mostly based on the

product and insufficient in measuring higher order skills. When the findings of these studies were taken into consideration, there has been a need to investigate whether teachers use the attainments of the course and re-review the functionality of it use in practice.

The majority of the students thought that the content of the course was sufficient. They were only indecisive in statistical and mathematical subjects. It can also be said that they were even questioning the necessity of these subjects. It can be interpreted that this questioning was arising from mathematical problems because most of the students stated that they had difficulty in statistical calculations, which was because they felt the lack of mathematical and statistical knowledge and skills. Similar to students, lecturers mentioned that they had difficulty in teaching the unit of statistics. Most of the participants who were students in English language department stayed away from mathematics because they studied language or verbal disciplines in high schools. This negatively affected teaching-learning processes of courses which required mathematical calculations. Many of the students suggested that Measurement and Evaluation course should contain examples and measurement instruments related to reading, writing, speaking and listening skills in English. Lecturers expressed similar opinions on this issue; therefore, a subject about teaching measurement instruments peculiar to the subject area and examples for these instruments could be added to the content of the course. Campbell and Evans (2000) also pointed out that the preservice teachers, who have successfully completed the measurement and evaluation course including the knowledge and skills regarding their own fields, may have an understanding about general evaluation practices by gaining compulsory knowledge in measurement and evaluation.

According to students' views, such activities as lecture, discussion, question-answer, examples, and exercises were used in the teaching process of the course; that is, mostly teacher-centered and partly student-centered methods and techniques were utilized and students pointed out that all these were sufficient for the course. In parallel with this, lecturers stated that they used the methods and techniques mentioned by students. Students remarked that they would like to make more practice, to see more examples, to respond more questions and to complete their projects step by step with their lecturers. In this regard, lecturers proposed that such practices as generating questions and developing tests should be focused more. Regarding this, similar studies highlighted that more practice should take place in the course (Birgin & Gürbüz, 2008; Çobanoğlu-Aktan & Çepni, 2010; Şahin & Abalı-Öztürk, 2014). Again, in the study conducted by Zhang and Burry-Stock (2003), it was determined that preservice teachers suggested that the evaluation projects should be integrated with teaching and other practical experiences. In addition, both lecturers and students stressed that the course content was too intense, so it could not be taught in one semester and package software or instructional technology should be used more in the course. In this respect, the course can be given throughout two semesters and classrooms can be organized as appropriate for group activities. Also, technological classrooms can be prepared, where there are statistical software programs that can automatically do calculations. In this way, problems deriving from students' lack of mathematical skills and knowledge can be resolved.

The great majority of the students can be said to find measurement-evaluation process of the course was sufficient. Whereas some students thought that process-based evaluation was more effective, some of them thought that product-based evaluation was more effective. Relatedly, students who found process-based evaluation more effective were more in number and there were students who suggested making various regulations in process-based evaluation. As for the views of lecturers on measurement-evaluation process of the course, it was generally identified that they applied identical measurement and evaluation practices; they had to use product-based evaluation because it was practical and easy to use and they had so many students. Also, they thought that only using product-based evaluation was not sufficient. Similar to these findings, Volante and Fazio (2007) found out that pre-service teachers tended to come into prominence summative evaluation rather than formative evaluation. Again in terms of students' readiness, both students and lecturers lacked of mathematical knowledge and skills and they had negative attitudes toward mathematics. They were also found out to have prejudice and concern for the course. This situation may create obstacles for gaining attainments of the course in an effective way. Hence, various precautions can be developed in order to erase negative attitudes and lack of knowledge and skills about mathematics. In relation to measurement-evaluation process of the course, students put forward different views although lecturers emphasized evaluation of skills and practices.

Besides, lecturers were in the opinion that teaching the course in the same way in all departments was wrong because they had difficulty in finding examples about the subject area and measurement and evaluation techniques would be different when students became real teachers. For this reason, it was revealed that lecturers needed to organize Measurement and Evaluation course in collaboration with other departments. Additionally, lecturers were found to use a syllabus instead of a curriculum. Lastly, lecturers stated that the existing practices should be revised again rather than constructing a new curriculum. In the study conducted by Özberk and Ünsal-Özberk (2012), it was indicated that there was a need to develop a curriculum in accordance with views of experts. That is to say, it can be suggested developing an up-to-date curriculum and revising Measurement and Evaluation course by cooperating with the departments.

References

- Arter, J. (1999). Teaching about performance assessment. *Educational Measurement: Issues and Practice*, 18 (2), 30-44. <https://doi.org/10.1111/j.1745-3992.1999.tb00012.x>
- Bachor, D. G., & Baer, M. R. (2001). An examination of pre-service teachers' simulated classroom assessment practices. *The Alberta Journal of Educational Research*, XLVII(3), 244-258.
- Birgin, O. (2006). The problems encountered in the application of portfolio evaluation method in primary education and the solution suggestions. Paper presented at the *First Student Symposium on Mathematics Education* (p. 39). İzmir, Turkey, Dokuz Eylül University.
- Birgin, O., & Gürbüz, R. (2008). Eliciting pre-service primary teachers' knowledge level about measurement and assessment. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 20, 163-179.
- Brookhart, S. M. (1999). Teaching about communicating assessment results and grading. *Educational Measurement: Issues and Practice*, 18(1), 5-13. <https://doi.org/10.1111/j.1745-3992.1999.tb00002.x>
- Çakan, M. (2004). Comparison of elementary and secondary school teachers in terms of their assessment practices and perceptions toward their qualification levels. *Ankara Üniversitesi Eğitim Fakültesi Dergisi*, 37(2), 99-114.
- Campbell, C., & Evans, J. A. (2000) Investigation of pre-service teachers' classroom assessment practices during student teaching. *The Journal of Educational Research*, 93(6), 350-355. <https://doi.org/10.1080/00220670009598729>
- Çobanoğlu, A. D., & Çepni, Z. (2010). Expert and teacher opinions on the scope and requirements of the measurement and evaluation course: pilot study. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(2), 85-92.
- Cutsforth, J. (2013). Examining pre-service teachers' understandings about classroom assessment. *Ohio Social Studies Review*, 50(2), 52-62.
- Devers, K. J., & Frankel, R. M. (2000). Study design in qualitative research–2: sampling and data collection strategies. *Education for Health*, 13(2), 263–271. <https://doi.org/10.1080/13576280050074543>
- Duban, N., & Küçükylmaz, E. A. (2008). Primary education pre-service teachers' opinions regarding to the use of alternative measurement-evaluation methods and techniques in practice schools. *İlköğretim Online*, 7(3), 769-784.
- Erdal, H. (2007). *2005 İlköğretim matematik programı ölçme değerlendirme kısmının incelenmesi (Afyonkarahisar ili örneği)* [The investigation of measurement & evaluation parts in the new elementary school mathematics curriculum (case of Afyonkarahisar)] (Unpublished master's thesis). Afyon Kocatepe University, Afyonkarahisar, Turkey.
- Freebody, P. (2004). *Qualitative research in education – interaction and practice* (2nd ed.). London: Sage Publications.
- Gelbal, S., & Kelecioğlu, H. (2007). Teachers' proficiency perceptions of about the measurement and evaluation techniques and the problems they confront. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- Gullickson, A. R. (1987). Teacher education and teacher-perceived needs in educational measurement and evaluation. *Journal of Educational Measurement*, 23(4), 347-354. <https://doi.org/10.1111/j.1745-3984.1986.tb00254.x>
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309-333. https://doi.org/10.1207/S15324818AME1503_5
- Hoover, N. R., & Abrams, L. M. (2013). Teachers' instructional use of summative student assessment data. *Applied Measurement in Education*, 26(3), 219-231. <https://doi.org/10.1080/08957347.2013.793187>
- Kilmen, S., & Çıkrıkçı-Demirtaşı, N. (2009). The perceptions of primary school teachers about their application levels of measurement and evaluation principles. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 42(2), 027-054. https://doi.org/10.1501/Egifak_0000001175
- Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice Oxford University Press*, 13(6), 522-525. <https://doi.org/10.1093/famp/13.6.522>
- 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. <https://doi.org/10.1111/j.1745-3992.2003.tb00142.x>
- Merriam, S. B. (1995). What can you tell from an N of 1: issues of validity and reliability in qualitative research. *PAACE Journal of Lifelong Learning*, 4, 51-60.

- Mertler, C. A. (2005). The role of classroom experience in pre-service and in-service teachers' assessment literacy. *Mid-Westerns Educational Researcher*, 1(4), 25-34.
- Özberk, E. H., & Ünsal-Özberk, E. B. (2012). Program development study based on need analysis in higher education: measurement and evaluation course model. *Eğitim ve Öğretim Araştırmaları Dergisi*, 1(4), 216-225.
- Özsevgeç, T., Çepni, S., & Demircioğlu, G. (2004). Measurement and evaluation literacy levels of science teachers'. Paper presented at the *Sixth Congress on National Science and Mathematics Education*, İstanbul, Turkey, Marmara University.
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). London: Sage Publications.
- Punch, K. F. (2005). Introduction to social research: quantitative and qualitative approaches. Ankara: Siyasal Kitabevi.
- Rodriguez, M. C. (2004) The Role of Classroom Assessment in Student Performance on TIMSS. *Applied Measurement in Education*, 17(1), 1-24, DOI: 10.1207/s15324818ame1701_1
- Şahin, Ç., & Abalı, Ö. Y. (2014). Opinions of prospective teachers on alternative assessment-evaluation methods. *Kastamonu Eğitim Dergisi*, 22(1), 123-142.
- 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. <https://doi.org/10.2307/20466661>
- Yıldırım, A., & Şimşek, H. (2008). *Qualitative Research Methods in Social Sciences*. Ankara: Seçkin Yayıncılık.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education*, 16(4), 323-342. https://doi.org/10.1207/S15324818AME1604_4

Copyrights

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

This is an open-access article distributed under the terms and conditions of the [Creative Commons Attribution license](#) which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.