NEVŞEHİR HACI BEKTAŞ VELİ ÜNİVERSİTESİ
Ara, bul.

Kapadokya Eğitim Dergisi
Cappadocia Journal of Education
Cilt: 6 Sayı: 1 Haziran 2025
Volume: 6 Issue: 1 June 2025
E-ISSN: 2717-9346

CHALLENGES FACED BY SCIENCE TEACHERS IN ASSESSING SKILLS IN STEM EDUCATION[1]

FEN BİLİMLERİ ÖĞRETMENLERİNİN STEM EĞİTİMİNDE BECERİ ÖLÇMEDE YAŞADIKLARI PROBLEMLER

Serdar VARİNLİOĞLU
Millî Eğitim Bakanlığı, Fen Bilimleri Öğretmeni
serdarv38@gmail.com.tr
ORCID: 0000-0002-4025-2419

Oktay BEKTAŞ
Erciyes Üniversitesi, Eğitim Fakültesi, Fen Bilgisi Eğitimi
obektas@erciyes.edu.tr
ORCID: 0000-0002-2562-2864.

Geliş/Received:
10.01.2025
Kabul/Accepted:
18.04.2025
Yayın/Published:
30.06.2025
Keywords
Science education,
STEM,
Skill measurement,
Phenomenology
Anahtar Kelimeler
Fen eğitimi,
STEM,
Beceri ölçme,
Fenomenoloji

ABSTRACT

This study examined the challenges that science teachers encounter when assessing students' skills in STEM tasks. The research employed a phenomenological design consistent with the attributes of qualitative research. Teachers who incorporate STEM activities in scientific curricula are favored. The research study group comprises seven science teachers employed in K*** and N*** provinces for the 2021-2022 academic year. Semi-structured interviews comprising two open-ended questions were executed. The first researcher performed online interviews, obtaining audio recordings with the participants' consent. Following the transcription of the audio recordings, a content analysis was performed by establishing categories, subcategories, and codes. The study found that science teachers encounter multiple challenges regarding accomplishment, time, student engagement, active participation, and integration of STEM disciplines while assessing skills in STEM activities. These issues include the lack of alternative assessment tools, students' difficulty in addressing skill-based questions, their familiarity with conventional assessment methods, insufficient preparedness, the challenge of measuring integrated skills, and overcrowded classrooms.

ÖZ

Bu çalışmada fen bilimleri öğretmenlerinin STEM etkinliklerini gerçekleştirme esnasında beceri ölçme açısından yaşadıkları sorunları araştırılmıştır. Çalışmada, nitel araştırma yönteminin özelliklerine uygun olarak, *fenomenoloji deseni* kullanılmıştır. Fen bilimleri dersinde STEM etkinliklerini uygulayan öğretmenler tercih edilmiştir. Araştırmanın çalışma grubunu 2021-2022 eğitim-öğretim yılında K*** ve N*** illerinde görev yapmakta olan yedi fen bilimleri öğretmeni oluşturmaktadır. İki açık uçlu sorudan oluşan yarı yapılandırılmış görüşmeler gerçekleştirilmiştir. Görüşmeler birinci araştırmacı tarafından çevrimiçi olarak gerçekleştirilmiş ve katılımcılardan izin alınarak ses kaydı alınmıştır. Ses kayıtları yazıya döküldükten sonra kategoriler, alt kategoriler ve kodlar oluşturularak içerik analizi yapılmıştır. Çalışma sonucunda STEM etkinliklerinde beceri ölçerken fen bilimleri öğretmenlerinin kazanım, süre, öğrenci, aktif katılım ve STEM disiplinlerinin entegrasyonu açılarından çeşitli problemlere sahip olduğu ortaya çıkmıştır. Bu problemler arasında; alternatif ölçme aracı bulamama, öğrencilerin beceri sorularını yetiştirememesi, öğrencilerin geleneksel ölçme araçlarına daha alışkın olması, düşük hazır bulunuşluk, entegrasyonda beceri ölçümü yapılamaması ve sınıfların kalabalık olması vardır.

DOI: *https://doi.org/10.69643/kaped.1622662*

Atıf/Cite as: Varinlioğlu, S. & Bektaş, O. (2025). Challenges faced by science teachers in assessing skills in STEM education. *Kapadokya Eğitim Dergisi, 6*(1), 97-114.

[1] Bu çalışma, birinci yazarın ikinci yazar danışmanlığında yürüttüğü doktora tezinden üretilmiştir.

Introduction

Constructivism emphasizes formative assessment, in which the teacher observes and evaluates as students apply and internalize new knowledge during the learning process (Maslovaty & Kuzi, 2002; Oakes & Lipton, 1999). Therefore, the constructivist learning approach requires various types of assessment that more deeply examine the structure and quality of students' learning and understanding (Hickey & McCaslin, 2001; Vadeboncoeur, 1997). Significant developments in the nature and conceptualization of assessment and evaluation have been observed in recent years, driven by the growing importance of constructivism (Gipps, 1999). Assessment and evaluation have evolved into both the outcome of the teaching-learning process and an integral component of it (Brooks & Brooks, 1992). In this context, assessment and evaluation have experienced a paradigm shift from a traditional to a broader educational assessment and evaluation approach, transitioning from a culture of tests and exams to a culture of assessment (Gipps, 1994; Murphy & Torrance, 1988).

The alternative assessment emerging from the new paradigm promotes higher-order thinking, encompasses learning outcomes and processes, integrates with instruction, and enables students to assess their work (Shepard, 2000). Traditional assessment tools generally determine the extent to which cognitive skills have been acquired. For this reason, the importance of measuring other skills is also increasing (Kylonen, 2012).

The constructivist approach has improved understanding of the relevance of students developing the previously mentioned skills. (Bybee, 2013). Grounded in the constructivist approach, STEM education is a learning methodology that promotes the development of these skills (Akgündüz, 2019). STEM education fosters a blend of knowledge and skills across various disciplines, enabling a thorough development and assessment of students (Honey et al., 2014; NRC, 2012). However, the problems encountered in measuring and evaluating students' skills during the STEM education process have not yet been sufficiently understood (Brophy et al., 2008). The issues faced in assessing the abilities obtained by students through STEM education and analyzing their underlying causes have been inadequately addressed (Buhagiar, 2007; Çepni, 2018).

Skill refers to the capacity to attain a particular outcome with optimal precision while minimizing the expenditure of energy and time (Guthrie, 1952). The world is continuously evolving, leading to the emergence of increasingly complex skills (WEF, 2020). The 2024 Maarif Model Program in Türkiye identifies skills comprising domain, conceptual, social-emotional learning, and literacy skills (MoNE, 2024). This program highlights the importance of imparting various skills, particularly vital skills such as 21st-century skills, to students. In STEM education, students must acquire problem-solving, creative thinking, decision-making, and critical thinking (Brophy et al., 2008). Honey et al. (2014) identify several goals of STEM education for students, including STEM literacy, the development of 21st-century skills, preparedness for the STEM workforce, fostering interest and engagement, and the capacity to integrate knowledge across STEM disciplines. It is as important to measure the extent to which students have acquired these skills as it is for them to acquire them themselves (Crane et al., 2003). In STEM education, assessment tools should be of a nature that measures skills. In other words, assessment in STEM education should be capable of measuring the higher-order thinking skills mentioned above (Fan & Yu, 2017; Saxton et al., 2014).

Although measuring skills in STEM education is important, there are issues in measuring various skills, especially 21st-century skills. Therefore, it is necessary to focus on what can be done to measure and evaluate these skills (Çepni, 2018). The literature suggests designs such as rubrics for measuring skills, but there are still limited quality assessment tools suitable for STEM education (Saxton et al., 2014).

There are two characteristics that the assessment and evaluation of quality STEM education should possess (Potter, 2017; Saxton et al., 2014; Sondergeld, 2014; Srinivasan, 2015). Firstly, assessment in STEM education should enable students to demonstrate their knowledge and skills. Secondly, assessment in STEM education should be able to measure various skills. Therefore, assessment in STEM education should be able to identify higher-order thinking skills such as cognitive, affective, psychomotor, and 21st-century skills (Harwell et al., 2015; Potter, 2017; Saxton et al., 2014; Sondergeld, 2014; Srinivasan, 2015; Tan & Leong, 2014; The Dayton Regional STEM Center, 2017). Teachers implementing STEM activities should perform assessments and evaluations based on the specified characteristics. Therefore, it is not sufficient for these teachers to use traditional assessment tools in STEM education (Odabaşı, 2018; Şardağ et al., 2018). Utilizing alternative assessment tools in conjunction with traditional assessment tools presents particular challenges for educators.

In this context, this study aims to determine the opinions of science teachers who implement STEM activities regarding the problems they encounter while using alternative assessment tools.

The literature reveals the inadequacy of teachers in alternative assessment (Tao, 2020). Therefore, studies have shown that teachers are not sufficiently competent in assessing STEM activities (Fan, 2024; Zengin et al., 2020). Some reasons show why teachers cannot use alternative assessment tools to measure various skills. For example, the cost of photocopying, difficulties in preparing and organizing alternative assessment tools (Gözütok et al., 2005; Şekel, 2007), difficulties experienced during implementation, parental reluctance (Bayraktar & Çınar, 2010), the central examination system's emphasis on traditional assessment tools, students' familiarity with results-oriented examination systems (Tezcan Şirin et al., 2022), the time-consuming nature of evaluating alternative assessment tools, the inability to evaluate alternative assessment tools objectively, issues related to validity and reliability, the inability to manage the classroom during the implementation of alternative assessment tools (Büyüktokatlı & Bayraktar, 2014), and the inadequacy of classroom physical structures (Radloff & Guzey, 2017). Therefore, this study has identified the problems teachers face implementing STEM activities when measuring their students' skills using alternative assessment tools. Thus, the authors identified problems similar to or different from those mentioned in the literature and aimed to shed light on educators who wish to address the issues in alternative assessment and evaluation. In this context, the study will contribute to the literature and raise awareness among researchers.

The lack of studies on alternative assessment in STEM education (e.g., Radloff & Guzey, 2017) has resulted in teachers implementing STEM activities possessing limited awareness of assessment. Therefore, more studies should be conducted on alternative assessment tools to solve the problems teachers face (Jeong et al., 2020). Although there are many assessment tools in the literature to measure STEM-related skills, finding and implementing all of these tools takes a significant amount of time for teachers (Çil & Çepni, 2018). Therefore, there is a need for comprehensive and multidimensional alternative assessment tools that will create a culture of measurement and evaluation. To rectify this shortcoming, it is necessary to identify problems when measuring skills. From this perspective, this study will constitute an important step in initiating positive actions to measure skills in future STEM activities.

Another notable point in the literature is that there is minimal emphasis on alternative assessment and evaluation methods and their issues in studies related to the implementation of STEM activities (Sarıcan, 2017). Therefore, there are very few studies on assessment and issues in STEM education (Zengin, 2021). From this perspective, this study will fill this gap in the literature and contribute to developing alternative assessment tools by identifying the problems of science teachers. As the reason for this situation, the authors of this study argue that alternative assessment tools should be used to determine the extent to which students have acquired skills in STEM activities. Thus, by measuring more skills, an assessment suitable for STEM education is conducted (Kutlu, 2006). Therefore, the authors of this study argue that the primary focus of the alternative assessment process used by teachers to measure their student's skills should be on identifying the problems. When these problems are identified, they believe practical alternative assessment tools can be used in STEM education. From this perspective, based on their opinions, this study aims to identify the problems experienced by science teachers who implement STEM activities in their classrooms while measuring skills in STEM education. In line with this purpose, the research question of this study is: What are the problems faced by science teachers who measure their students' skills in STEM education?

Method

Research Design

This study has utilized phenomenology, which is the design of qualitative research. Phenomenology is a research design that examines events, perceptions, and experiences recognized but not fully understood in detail (Patton, 2018). This research employs a phenomenological design to thoroughly examine science teachers' challenges concerning skill measurement in STEM activities, which they recognize but do not fully comprehend.

Study Group

This study used criterion and maximum diversity sampling, which are types of purposive sampling, by varying the sample. Criterion sampling involves the examination of all instances that satisfy established criteria

(Patton, 2018). This study selected science teachers who implement STEM activities in science classes as the study group based on the criterion. Establishing a maximum diversity study group aims to identify common occurrences across varied contexts (Creswell, 2013). This research sought to identify common occurrences in STEM practices related to gender and school context by examining the participants' gender and their respective schools. Therefore, the study group was formed by selecting two women and five men, seven science teachers working in various public and private middle schools, science art center (BİLSEM), and public high schools affiliated with the Ministry of National Education (MoNE) in the provinces of K*** and N*** during the 2021-2022 academic year. The study group selection considered the voluntary participation of science teachers and their instruction of science across various grade levels. The participants have the pseudonyms Cankat, Harun, Sevim, Buse, Murat, Ersin, and Mehmet. The demographic information of the science teachers in the study group is provided in Table 1.

Table 1. Demographic Information of the Participants

Participants	Gender	The classes s/he teaches	The school where he/she is assigned
Cankat	Male	Fifth-eighth grade	Science art center
Ersin	Male	Sixth grade	The state middle school
Buse	Female	Seventh grade	The state middle school
Sevim	Female	Ninth grade	The state high school
Harun	Male	Fifth grade	The state middle school
Mehmet	Male	Fifth grade	The state middle school
Murat	Male	Fifth grade	Private middle school

Data Collection Tools

This study used semi-structured interviews. An interview is an interactive data collection tool conducted in a question-and-answer format for a predetermined, planned, and serious purpose (Creswell, 2013; Stewart & Cash, 1985). In this study, one of the reasons for using interviews is the researchers' desire to gain a more detailed understanding of the issues faced by science teachers in measuring skills during the implementation of STEM activities. Additionally, the high response rate of participants during the interview and the ability to record their immediate reactions to questions can be cited as another reason (Creswell, 2013).

Data Collection Process

The necessary institutional permissions and ethics committee approval (No. 105) for this study have been obtained. Before starting the interview process, a literature review was conducted to create a pool of interview questions. The opinions of a science educator were sought on which of these questions would be included in the research. Therefore, questions have been created to identify the problems that science teachers encounter while measuring skills during the implementation of STEM activities. In this context, an interview form has been prepared. The interview form consists of sections that include general information-ethical explanations, personal information, and open-ended questions. There are two open-ended questions and probes in the interview. The interviews were conducted online via Zoom by the first researcher, lasting approximately 30-45 minutes. Video and audio recordings were made with the participants' consent during the interview, provided that ethical rules were followed. The audio recordings obtained during the interview were transcribed and put into written form. After the interview, the recorded discussions were transcribed and confirmed by the participants, who signed them.

Data Analysis

This study used content analysis. Content analysis is a type of data analysis that aims to reach concepts that can explain the obtained data and the relationships between these concepts (Merriam, 2018). The reason for using content analysis in this study is to reach concepts and inter-conceptual relationships related to the difficulties experienced in measuring skills. To address this issue, we initially examined the sections that lacked clarity due to the interview data being recorded in spoken language. In cases where a word is reiterated in spoken language, it is documented singularly in written form. The places where the participant made semantic

errors have been attempted to be corrected. Instead of the small, meaningless conversations the participants asked themselves, ellipses (…) have been used.

Secondly, categories, subcategories, and codes have been created to identify meaningful data sets. In this process, the researchers identified two different categories. Ideas were gathered from the supervisor regarding the categories and codes, and necessary corrections were made based on the feedback received. For example, the expert suggested that instead of the code "unable to find an alternative assessment tool suitable for the subject-learning outcome in measuring cognitive skills," it might be better to revise it to "unable to find an alternative assessment tool." After the necessary reductions, the researchers reached a consensus and concluded the analysis process. These categories and subcategories are presented in Table 2. The findings section presents each category in detail under its respective heading.

Table 2. Categories and Subcategories

Categories	Subcategories
Problems in Measuring Skills in STEM	Objectives Duration Student level Active participation Integration of mathematics, technology, and engineering Number of students
Unmeasurable Skills in STEM and Their Justification	-

Validity and Reliability

Validity is categorized into two types: internal validity and external validity (Merriam, 2018). This research utilizes the findings of Eroğlu and Bektaş (2016) as a study of internal and external validity. Table 3 delineates the validity assessments performed by the authors.

Table 3. Actions Taken for Validity Control

Validity	Internal validity	Long-term interaction Participant and environment confirmation Sample triangulation Expert review Direct quote
	External validity	Purposeful sampling choice Detailed descriptions The role of the researcher

The first researcher created a communication group on their mobile phone for the interview. Before the interview, the researcher gave a calming talk to reduce the participant's excitement. The first researcher read the purpose of the study before starting the interview, obtained permission for the interview recording, and stated that the information about the participants would remain confidential. The researcher avoided words and actions that would disturb the participant during the interview. Additionally, he refrained from asking leading questions or exhibiting guiding behaviors toward the participant. The interviews lasted between 30-45 minutes. After the participant's responses, the first researcher summarized the answers to the participant and expected the participant to confirm the responses. The first researcher ensured that interview was conducted in a quiet environment with uninterrupted internet connectivity. After the interview, the transcripts were signed by the participants. The study employed criterion and maximum diversity sampling methods to establish the study group. Therefore, a variety of study groups have been formed. Within the scope of the expert review, the authors sought the opinions of an expert science educator on the selection of data collection tools, the preparation and finalization of the interview form, and the organization of codes and categories in qualitative research. The authors provided direct quotations related to the codes in the findings.

The authors performed three distinct studies, detailed in Table 3, to validate the transferability of the research for other researchers. Initially, science teachers who engaged in the implementation of STEM activities were selected. Participants for the study group were selected based on criteria including gender, grade level, and the

schools in which they were employed. Secondly, each section of the study has been documented in comprehensive detail. The role of each researcher has been clearly defined throughout the study.

Table 4. Actions Taken for Reliability Control

Reliability	Internal reliability	Consensus among coders Presenting findings without commenting on them
	External reliability	Verifying the alignment among the data analysis, results, and discussion sections with a specialist in science education

Table 4 shows the researchers' internal and external reliability studies. In internal reliability studies, the researchers reached a consensus on the codes and presented the findings without interpretation. As an external validity study, the researchers had the data analysis, results, and discussion sections verified by an expert in qualitative research in science education.

Results

The researchers categorized the findings into two distinct groups. The authors have presented the codes for each subcategory in tables. Direct quotes were employed as evidence to support each code. The authors have incorporated direct quotes from various participants. The writers chose to include just one quote from each of the three frequency groups (high, medium, and low).

Problems in Measuring Skills in STEM

Objectives

Table 5 presents the codes associated with participants' perceptions of the challenges encountered in assessing skills in STEM relative to objectives. The participants recognized issues categorized under six distinct codes. The writers have provided direct quotations corresponding to each code in Table 5.

Table 5. Codes for Participants in the Objectives

Participants	Codes
Mehmet, Ersin, Murat, Harun	Inability to find an alternative assessment tool
Harun, Ersin, Murat	The inability of teachers to develop assessment tools appropriate to the learning outcomes
Mehmet, Murat	Students' grade anxiety
Mehmet, Harun	The incompatibility of alternative assessment tools with the grading system
Murat	The inability to measure higher-order thinking skills, students giving random answers to assessment tools

Mehmet, Ersin, Murat, and Harun expressed the difficulty of locating alternative assessment instruments for measuring cognitive skills. Mehmet asserts, *"Challenges arise in identifying assessment forms relevant to biology objectives. Assessment tools related to environmental subjects, cellular division, DNA, mutations, alterations, organisms, energy, photosynthesis, and respiration are currently limited in availability. I believe it is an issue."* Harun, Ersin, and Murat contend that teachers cannot develop assessment instruments that align with the learning outcomes. Ersin asserts, *"I encountered challenges in utilizing (developing) alternative assessment tools while executing biology-related activities due to my difficulties in performing the tasks."* Murat raised the concern over the inability to assess higher-order thinking skills. Murat states: *"... our challenge lies in integrating objective into daily life...The primary issue is present. He understands the concept but faces considerable difficulties in its practical application. A challenge necessitating exploration and analysis... We are struggling with everyday life problems."*

Duration

Table 6 shows the codes related to the participants' thoughts on the problems experienced in measuring skills in STEM in terms of time. The participants identified issues categorized under five distinct codes. Direct quotes related to each code are presented in Table 6.

Table 6. Codes for Participants in the Duration

Participants	Codes
Murat	Students' inability to complete skill questions, the time it takes to measure abstract concepts
Ersin	The teachers do not have enough time to prepare the guidelines for the assessment tools
Cankat	The measurement of affective skills takes time.
Harun	The process of obtaining assessment tools suitable for measuring cognitive skills is time-consuming.

The authors provided direct quotes from Murat, who has two different codes in the duration category, and Cankat, who has one code. Indicating that he experienced problems in measuring time, Murat expressed this issue from both the student and teacher perspectives. For example, concerning the challenge faced by students in completing skill-based questions, Murat thought: *"When a child moves from seventh grade to eighth and sees the skill questions related to adapting the objectives learned under the new generation to daily life, they experience much difficulty. The time is not enough."* Cankat, on the other hand, stated that he had trouble assessing the students' affective traits. Cankat stated: *"...some affective gains are not gains that can be measured in a short time. A study should be conducted for at least eight weeks to measure an attitude effectively. The literature says so. Therefore, if you say we should constantly measure affective skills in real-time, I think you would be mistaken."*

Student Level

Table 7 presents the codes associated with the participants' reflections on the challenges faced at the student level in assessing skills in STEM. The participants identified issues categorized under nine distinct codes. Table 7 presents direct quotes associated with each code.

Table 7. Codes for Participants in the Student Level

Participants	Codes
Murat, Sevim, Ersin, Harun	Students' familiarity with traditional assessment tools
Murat, Mehmet, Harun	Low readiness
Mehmet	The lack of STEM infrastructure for students
Harun	Students receive help from their surroundings.
Ersin	The relegation of affective skills to the background
Cankat	Students' biases towards skill-related assessment forms, Students' reluctance to have process-based assessments take a long time
Murat	The lower grades experience adaptation problems in the assessment and evaluation process, and the lack of teaching problem-solving skills to students in the lower grades.

Murat, Sevim, Harun, and Ersin have talked about the issue of students being accustomed to traditional assessment tools. Harun stated: *"Due to the high school entrance exam (LGS), when we ask the students to do an activity, they say, let us solve multiple-choice questions. The students have focused on the question. The child sees these activities as a waste of time. Let us solve questions, Why should we bother with these? (they say). I tend to be more hesitant in measuring various skills in eighth graders."* Ersin mentioned the issue of affective skills being sidelined in assessments. Ersin explained: *"I can say that I felt more comfortable measuring cognitive and psychomotor skills in the fifth, sixth, and seventh grades. However, I think that the anxiety of the LGS negatively affects the measurement of these skills in eighth grade as well. Due to the exams, measuring affective skills is being sidelined."*

Active Participation

Table 8 presents the codes associated with the participants' reflections on the challenges faced in actively engaging in measuring skills within STEM. The participants have categorized the problems into four distinct codes. Table 8 presents direct quotations associated with each code.

Table 8. Codes for Participants in the Active Participation

Participants	Codes
Mehmet, Cankat, Sevim	Low readiness
Murat, Cankat	Students' prejudice
Mehmet	The inadequacy of physical conditions, The lack of STEM infrastructure for students

Mehmet, Cankat, and Sevim have mentioned the issue of low preparedness. Sevim stated: *"Unfortunately, solving problems is their top priority for some students. Therefore, the activities you do to develop and measure their creativity have little impact on those students."* Murat and Cankat have mentioned the issue of students' biases. Murat explained: *"There is a problem with active participation. If you want, use the best of the Web 2.0 tools. If the youngster is prejudiced, they perceive assessment and evaluation unfavorably."*

Integration of Mathematics, Technology, and Engineering

Table 9 presents the codes associated with participants' perceptions of the challenges encountered in integrating mathematics, technology, and engineering in assessing skills within STEM. The participants have delineated issues categorized under six distinct classifications. Table 9 presents direct quotations corresponding to each code.

Table 9. Codes for Participants in the Integration of Mathematics, Technology, and Engineering

Participants	Codes
Ersin, Mehmet, Cankat, Sevim	The inability to measure skills in the integration of science with mathematics and technology
Murat	The inadequacy of mathematics achievements
Ersin	The centrality of science and engineering
Mehmet	The lack of integration and assessment knowledge among teachers, Only the measurement of science subjects, Insufficient physical conditions

Ersin, Mehmet, Cankat, and Sevim have identified the challenge of assessing competencies in integrating science with mathematics and technology. For example, Sevim stated: *"... mathematical modeling is the area where we are weakest because they have almost completely removed mathematics from the chemistry content. We do not rely heavily on calculations. I make sure that students do mathematical calculations as much as possible"*. Ersin mentioned the issue of science and engineering being at the center. Ersin stated: *"I can say that I can measure the integration of science and engineering more easily because I can say that the children produced a product and that these products emerged as a result of engineering integration. I can say that I have a problem with measuring the integration of science, mathematics, and technology with alternative assessment tools."*

Number of Students

Table 10 presents the codes associated with participants' perceptions of the challenges encountered in assessing skills in STEM, quantified by the number of students. The participants have categorized the issues under a unified code. Table 10 presents direct quotes related to this code.

Table 10. Codes for Participants in the Number of Students

Participants	Code
Mehmet, Ersin, Harun	The overcrowding of the classrooms

Mehmet, Ersin, and Harun have mentioned the issue of crowded classrooms in skill assessment. Ersin stated: "*... As the number of students increases, our difficulties in cognitive, psychomotor, and affective assessments also increase.*"

Unmeasurable Skills in STEM and Their Justification

Participants stated that they could not measure logical reasoning, affective, psychomotor, 21st-century, mathematical modeling, interdisciplinary connection-making, design-oriented, and scientific process skills, and

they provided their justifications for this situation. Table 11 consists of codes that express the justifications provided by the participants for at least four skills. Table 11 presents direct quotations related to each code.

Table 11. Codes for Participants in the Unmeasurable Skills in STEM and Their Justification

Participants	Codes
Murat, Mehmet, Ersin, Harun, Sevim	Emphasis on cognitive skills
Mehmet, Cankat, Harun, Sevim	Time constraint
Ersin, Murat, Cankat, Harun	The teacher's inadequacy in assessment
Murat, Ersin, Cankat, Harun	The lack of sufficient alternative assessment tools and the difficulty in preparing them
Ersin, Murat, Mehmet	The inadequacy of physical conditions
Murat, Ersin, Harun	The intensive use of traditional assessment tools
Mehmet, Sevim	The inadequacy of the student-level
Ersin	The neglect of higher-order thinking skills in the scales
Cankat	The preference for scales in academic work, The lack of validity and reliability for the scales

Murat, Mehmet, Ersin, Harun, and Sevim expressed the reason for measuring various skills as the emphasis on cognitive skills. For example, Ersin stated: *"In our traditional assessment approach, cognitive skills are more prominent, and we use alternative assessment tools mainly to measure cognitive skills."* Murat, Ersin, Cankat, and Harun explained the reason for not measuring various skills as the lack of sufficient alternative assessment tools and the difficulty in preparing them. Harun declared, *"If we were provided with a lesson plan, sample scales, or materials to measure the skills, I believe we could easily handle it."* Cankat stated the reason for not measuring various skills as the preference for scales in academic work and the lack of validity and reliability studies. Cankat said: *"Validity and reliability studies need to be conducted. There may be many studies and theses, but it is debatable how accurate it is when applied to our BİLSEM students. Therefore, I have not used such a scale until now, but I would probably use it if we were to conduct a scientific study,"*

Conclusion, Discussion, and Recommendations

Problems in Measuring Skills in STEM

This research has concluded that science teachers face many problems when measuring skills in STEM activities. Teachers have problems regarding achievement, duration, students, active participation, and integration. Among these problems are the inability to find alternative assessment tools, students' inability to keep up with skill questions, students' familiarity with traditional assessment tools, low readiness, the inability to measure skills in integrating science with mathematics and technology, and overcrowded classrooms.

The authors have concluded that science teachers could not find alternative assessment tools to understand whether students have achieved STEM gains. This result is consistent with the studies of El Nagdi and Roehrig (2022), Anıl and Acar (2008), and Gerek (2006) in the literature. El Nagdi and Roehrig (2022) and Gerek (2006) stated that teachers favor traditional assessment tools for measuring and evaluating student skills. Gerek (2006) stated that the reason for teachers' preference for traditional assessment tools is that they are easy to prepare and evaluate. Anıl and Acar (2008) stated that teachers face serious difficulties using alternative assessment tools; they are not sufficiently informed about them, their implementation requires excessive time and financial resources, and they encounter problems preserving the assessment results. Therefore, the authors of this study claim that teachers prepare these tools themselves due to the lack of alternative assessment tools for STEM activities. This process leads teachers to abandon the assessment and evaluation process because they do not have sufficient knowledge to prepare alternative assessment tools. Therefore, the authors argue that teachers need training in preparing and implementing alternative assessment tools for measuring skills in STEM education.

The authors have determined that students cannot answer skill-based questions in STEM activities successfully. This outcome corresponds with the research conducted by Ünsal and Kaba (2022). Ünsal and Kaba (2022) asserted that skill-based questions enhance students' problem-solving and advanced cognitive skills. Conversely, Ünsal and Kaba (2022) asserted that students with low academic performance experience a sense of failure when they cannot resolve skill-based problems. The study's findings indicate that students'

familiarity with multiple-choice questions, inadequate reading comprehension abilities, and insufficient focus on assessing higher-order thinking skills in science classrooms may have contributed to their difficulties with skill-based questions. Consequently, the authors assert that science teachers should prioritize activities designed to cultivate higher-order thinking skills during STEM initiatives and allocate additional time to children with lower academic performance.

The authors have concluded that students are accustomed to traditional assessment tools in STEM activities. This result is consistent with Gelbal and Kelecioğlu's (2007) and Şenel Çoruhlu et al. (2009) studies in the literature. Gelbal and Kelecioğlu (2007) concluded that teachers faced problems implementing alternative assessment tools because these tools were not oriented toward centralized exams. The present study also attributes teachers' preference for traditional assessment tools over alternative assessment tools in STEM activities to multiple-choice questions used in centralized exams. Therefore, this study argues that alternative assessment tools should be used in STEM activities. This argument will raise awareness among those preparing to use alternative assessment tools for central exams.

The authors have concluded that students have a low level of preparedness in STEM activities, and this level negatively affects their active participation. This result is consistent with the study by Karakaya and Yılmaz (2022). According to Karakaya and Yılmaz (2022), the assessment and evaluation of STEM activities should be carried out with the active participation of teachers and students. Birzina et al. (2021) concluded that students are not ready for STEM learning and the evaluation process. Therefore, students' readiness is influenced by changes in learning objectives, knowledge, understanding, skills, habits, values, and attitudes (Kearney & Garfield, 2019). Therefore, science teachers should measure students' readiness in the classroom where they are conducting STEM activities for the first time and select activities and assessment tools accordingly. The current study argues that teachers who support students' active participation in line with the constructivist philosophy will determine their students' readiness. As a result, students can participate more effectively in STEM activities and achieve meaningful learning.

The authors have concluded that science teachers cannot measure their students' skills in integrating science with mathematics and technology in STEM activities. This result parallels the study by Gao et al. (2020). Gao et al. (2020) stated that few programs assess interdisciplinary skills. Similarly, Pimthong and Williams (2018) stated that pre-service teachers need interdisciplinary skill-based assessment and evaluation training. The authors believe that students' skills cannot be measured in integrating science with mathematics and technology due to STEM activities' multidimensional and complex nature. Additionally, the authors argue that the lack of adequate training for science teachers in measuring interdisciplinary skills is also a factor in their inability to conduct this assessment. Similarly, Gao et al. (2020) have emphasized that teachers evaluating students in STEM education should pay attention to interdisciplinary learning and practices. Gao et al. (2020) stated that the process of measuring interdisciplinary skills should be structured to provide feedback to students. Additionally, Gao et al. (2020) stated that evaluating interdisciplinary connections in student understanding would help determine whether an interdisciplinary STEM education program is achieving its intended goals. In conclusion, the current study's authors argue that measuring and evaluating interdisciplinary skills is important for effective STEM education and claim that this process will be improved through teacher training. Therefore, the authors argue that researchers should use research designs such as action research, experimental design, and nested design to provide teacher education.

The authors have concluded that science teachers face the problem of overcrowded classrooms during STEM activities. This result parallels the study by Radloff and Guzey (2017). Radloff and Guzey (2017) mentioned the importance of preparing a physical environment where the skills students will use to solve multidimensional problems in STEM activities can be evaluated. From this perspective, the authors of this study claim that teachers reduce the time they allocate for assessment and evaluation due to overcrowded classrooms. Additionally, the authors suggest emphasizing activities that involve group work to ensure that the time allocated for the assessment and evaluation process is sufficient for the students. Therefore, the authors argue that science teachers should use alternative assessment tools suitable for group evaluation.

Unmeasurable Skills in STEM and Their Justification

The research findings indicate that science teachers engaged in STEM activities cannot assess logical reasoning, affective, psychomotor, 21st-century, mathematical modeling, interdisciplinary connection-making,

design-oriented, and scientific process skills. Teachers face challenges in measuring these skills due to the focus on cognitive skills, insufficient alternative assessment tools, and the inadequacy of scales regarding validity and reliability. The authors concluded that science teachers prioritized cognitive skills as the primary factor contributing to their challenges in assessing various skills in STEM activities. Kaloyanova (2023) found that teachers predominantly assess cognitive skills in the context of STEM education and that each STEM activity possesses distinct characteristics, necessitating that educators assess various skills within each activity. This study argues that science teachers should evaluate students' diverse skills in STEM activities. Nevertheless, the authors contend that science teachers often prioritize assessing students' cognitive skills, as they are familiar with an educational framework that predominantly evaluates academic achievement. The authors suggest that an additional reason for this habit is the focus on assessing cognitive skill achievements within the science education program. This study contends that the middle school science curriculum in Türkiye must be structured to assess the skills pertinent to the STEM education process. This study claims that measuring these skills is unfeasible while the multiple-choice exam used for high school admission in Turkey remains in place.

The authors conclude that science teachers struggle to assess various skills in STEM activities because of insufficient alternative assessment tools and the challenges associated with their preparation. The findings align with the research conducted by Tekin Poyraz (2018), Saxton et al. (2014), and Zengin et al. (2020). According to Tekin Boyraz (2018), teachers face these challenges due to inadequate skills in assessment and evaluation, along with the time-intensive demands of developing alternative assessment and evaluation methods. Harris et al. (2023) emphasize the importance of teachers having sufficient alternative assessment tools for process-based assessments. Koştur (2023) asserts the necessity for advancements in the alternative assessment process within STEM education. Margot and Kettler (2019) assert the necessity for curricula that facilitate the implementation of alternative assessment tools. The authors of this study contend that science teachers face this issue due to inadequate literature review regarding alternative assessment tools. The inability of science teachers to access adequate resources through the English literature, the primary language of science, is a matter of concern. This study is significant for science teachers seeking to enhance their English proficiency.

The authors found that science teachers could not effectively use the scales to assess skills in STEM activities, as these scales were employed in academic research, and validity and reliability assessments could not be performed. This outcome aligns with the findings of Karakaya and Yılmaz (2022) who indicated that science teachers encountered difficulties due to their inability to design measurement tools with verified validity and reliability. The authors of this study contend that employing measurement tools that lack adequate validity and reliability assessments in STEM activities wastes time.

Suggestions

Based on findings:

- Science teachers and pre-service science educators require training in preparing and implementing diverse alternative assessment tools.
- Science teachers should use research designs such as action research, experimental design, and nested design in their study to improve the measurement of their students' skills.
- Science teachers should focus on activities that enhance their students' higher-order thinking skills in STEM events.
- Science teachers should give more time to students with low academic performance.
- Science teachers should use alternative assessment tools for group evaluation in crowded classrooms.
- Science teachers should assess students' readiness before starting STEM activities and accordingly select activities and assessment tools.
- The middle school science curriculum should be organized to measure the skills in the STEM education process.
- Science teachers should improve their foreign language skills to access alternative assessment tools in the literature.

Based on limitations:

- This study was conducted using the phenomenological design within the qualitative research method. Different studies addressing the issues encountered in skill assessment and evaluation in STEM activities can be conducted using the action research design within the qualitative research method.
- This study was conducted with science teachers implementing STEM activities in their classes. Similar studies can be conducted with pre-service science teachers
- In this study, a semi-structured interview was used as the data collection tool. Additionally, various studies can be conducted using different data collection tools.

Declaration of Contribution Rate of Researchers

This article is derived from the first author's doctoral dissertation. The researchers contributed equally to the planning, execution, and writing of this research.

Statement of Support and Acknowledgment

No support was received from any institution, organization, or person in this research.

Conflict Statement

The researchers do not have any financial or personal conflicts of interest with other institutions and individuals related to the research.

Ethics Committee Declaration

Permission for this study was granted by the Erciyes University Ethics Committee with the decision dated 29.03.2022 and numbered 105.

References

Akgündüz, D. (2019). Theoretical framework and historical development of STEM education. D. Akgündüz (Ed.), *STEM Education in Theory and Practice from Preschool to University (ss. 19-47)*. Anı Press.

Anıl, D. D., & Acar, M. (2008). Classroom teachers' views on the problems they encounter in the measurement and evaluation process. *Van Yüzüncü Yıl University Faculty of Education Journal, 5*(2), 44-61. https://dergipark.org.tr/tr/download/article-file/146332

Bayraktar, Ş., & Çınar, D. (2010). Levels of effective teacher behaviors of science and technology teachers through the eyes of prospective teachers. *Ahi Evran Üniversitesi Kırşehir Faculty of Education Journal*, *11*(3), 131-152. https://acikerisim.selcuk.edu.tr/server/api/core/bitstreams/a8ef41b7-81fe-439a-a49f-a29d23fa33fe/content

Birzina, R., Pigozne, T., & Cedere, D. (2021). Students' readiness for STEM learning within the context of national education reform. *Human, Technologies, and Quality of Education*, 657-752. https://doi.org/10.22364/htqe.2021

Brooks, J. G., & Brooks, M. G. (1992). *In search of understanding: the case for constructivist classrooms*.: Association for Supervision and Curriculum Development.

Brophy, S., Kleın, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *Journal of Engineering Education,* 97(3), 369–387. https://doi.org/10.1002/j.2168-9830.2008.tb00985.x

Buhagiar, M. A., (2007). Classroom assessment within the alternative assessment paradigm: Revisiting the territory. *Curriculum Journal, 18*(1), 39-56. https://doi.org/10.1080/09585170701292174

Büyüktokatlı, N., & Bayraktar, Ş. (2014). Alternative assessment practices in science. *Journal of Education and Training, 4*(1), 103-126. https://doi.org/10.14527/pegegog.2014.006

Bybee, R. W. (2013). *The case for STEM education: Challenges and opportunities.* NSTA Press.

Crane, T., Wilson, J., Maurizio, A., Bealkowski, S., Bruett, st K., & Couch, J. (2003). *Learning for the 21 century: A report st and mile guide for 21 century skills.* Retrieved November 1, 2024. from http://www.p21.org/images/stories/ other docs/p2 1up_Report.pdf

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage.

Çepni, S. (2018). *Kuramdan uygulamaya STEM eğitimi* (4. Baskı). Pegem Academy Press.

Çil, E., & Çepni, S. (2018). *Fen bilimleri dersi öğretim programı ilkokul ve ortaokul öğretmen el kitabı* (6. Baskı). Pegem Academy Press.

El Nagdi, M. A., & Roehrig, G. H. (2022). Reality vs. Expectations of assessment in STEM education: An exploratory case study of STEM schools in Egypt. *Education Sciences, 12*(11), 1–20. https://doi.org/10.3390/educsci12110762

Eroğlu, S., & Bektaş, O. (2016). Opinions of STEM-educated science teachers on STEM-based lesson activities. *Journal of Qualitative Research in Education, 4*(3), 43-67. https://doi.org/10.14689/issn.2148-2624.1.4c3s3m

Fan, S. C., & Yu, K. C. (2017). How an integrative stem curriculum can benefit students in engineering design practices. *International Journal of Technology & Design Education, 27*(1), 1–23. https://doi.org/10.1007/s10798-015-9328-x

Fan, Y. (2024). Integrating online and offline teaching to promote creativity for STEM learners. *Educational Technology Quarterly, 2024*(3), 241–254. https://doi.org/10.55056/etq.723

Gao, X., Li, P., Shen, J. & Sun, H. (2020). Reviewing assessment of student learning in interdisciplinary STEM education. *International Journal of STEM Education, 7*(24). https://doi.org/10.1186/s40594-020-00225-4

Gelbal, S., & Kelecioğlu, H. (2007). Teachers' perceptions of competence regarding measurement and evaluation methods and the problems they encounter. *Hacettepe University Faculty of Education Journal, 33*, 135-145. https://dergipark.org.tr/tr/download/article-file/87640

Gerek, Ö. (2006). *A study on classroom teachers' opinions, evaluations, and competencies about the new primary education curriculum.* [Master's thesis, Selçuk University]. National Thesis Center.

Gipps, C. V. (1994). *Beyond testing: towards a theory of educational assessment.* Routledge Falmer.

Gipps, C. (1999). Socio-cultural aspect of assessment. In A. Iran-Nejad & P.D. Pearson (Eds.), *Review of Research in Education*, 24 (pp. 355–392). American Educational Research Association.

Gözütok, D. (2005). *Evaluation of Primary Education Programs in Terms of Teacher Competencies*, Reflections in Education VIII: Symposium on Evaluating New Primary Education Programs, Tekışık Education Research and Development Foundation Publications, 17–40.

Guthrie, E. R. (1952). *The psychology of learning*. Harper.

Harris, C.J. Wiebe, E., Grover, S., & Pellegrino, J.W. (Eds.) (2023). *Classroom-Based STEM assessment: Contemporary issues and perspectives*. Community for Advancing Discovery Research in Education (CADRE). Education Development Center, Inc.

Harwell, M., Moreno, M., Phillips, A., Guzey, S. S., Moore, T. J., & Roehrig, G. H. (2015). A study of STEM assessments in engineering, science, and mathematics for elementary and middle school students. *School Science and Mathematics, 115*(2), 66-74. https://doi.org/10.1111/ssm.12105

Hickey, D. T. & Mc Caslin, M. (2001). A comparative, sociocultural analysis of context and motivation. In S. Volet & S. Jarvela (Eds.). *Motivation in learning contexts* (pp.33–56). Pergamon.

Honey, M., Pearson, G., & Schweingruber, A. (2014). *STEM integration in K-12 education: status, prospects, and an agenda for research*. National Academies Press.

Jeong, J. S., González-Gómez, D., & Yllana Prieto, F. (2020). Sustainable and flipped STEM education: formative assessment online ınterface for observing pre-service teachers' performance and motivation. *Education Sciences*, *10*(10), 283. https://doi.org/10.3390/educsci10100283

Kaloyanova K., Leventi N., Kaloyanova E. (2023). Evaluating computing students' digital skills and health literacy: A case from Bulgaria. *Front Public Health, 24*(10). https://doi.org/10.3389/fpubh.2022.1085842

Karakaya, F. & Yılmaz, M. (2022). Teachers' views on assessment and evaluation methods in STEM education: A science course example. *Journal of Pedagogical Research, 6*(2), 61-71. https://dx.doi.org/10.33902/JPR.202213526

Kearney, W. S. & Garfield, T. (2019). Student readiness to learn and teacher effectiveness: Two key factors in middle grades mathematics achievement, *RMLE Online, 42*(5), 1–12. https://doi.org/10.1080/19404476.2019.1607138

Koştur, H. İ. (2023). Predispositions define a pro-environmental attitude. *Journal of STEAM Education, 6*(1), 61-83. https://doi.org/10.55290/steam.1167600

Kutlu, Ö. (2006). Ways to identify higher-level mental processes: New situational assessment approaches, *Journal of Contemporary Education, 31*(335), 15–21. https://search.trdizin.gov.tr/tr/yayin/detay/74516

Kylonen, P. C. (2012, May 7-8.). *Measurement of 21st-century skills within the common core state standards*. Paper presented at the Invitational Research Symposium on Technology Enhanced Assessments.

Margot, K. C., & Kettler, T. (2019). Teachers' perception of STEM integration and education: A systematic literature review. *International Journal of STEM Education, 6*(2), 1–16. https://doi.org/10.1186/s40594-018-0151-2

Maslovaty, N., & Kuzi, E. (2002). Promoting motivational goals through alternative or traditional assessment. *Studies in Educational Evaluation, 28*, 199-222. https://doi.org/10.1016/S0191-491X(02)80002-5

MoNE (2024). Science Education Model Curriculum. MEB Basic Education General Directorate. https://tymm.meb.gov.tr/genel-bakis?id=4

Merriam, S. B. (2018). *Qualitative research* (3rd ed.). Nobel Press.

–

Murphy, R. & Torrance, H. (1988). *The changing face of educational assessment*. Milton Keynes, Open University Press.

NRC (National Research Council). (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas*. The National Academy Press.

Oakes, J., & Lipton, M. (1999). *Teaching to Change the World*. McGraw-Hill.

Odabaşı, Ş. Y. (2018). Hello STEM: An innovative teaching approach (1st ed.). In K. A. Kırkıç & E. Aydın (Eds.), *Assessment and evaluation for STEM* (pp.109–122). Education Press.

Patton, M. Q. (2018). *Nitel araştırma ve değerlendirme yöntemleri*, (Bütün, M., Demir, S.B. Çev.). Pegem Akademi Press.

Pimthong, P., & Williams, J. (2018). Preservice teachers' understanding of STEM education, *Kasetsart Journal of Social Sciences*, *41,* 289–295. https://doi.org/10.1016/j.kjss.2018.07.017

Potter, B. S., Ernst, J. V., & Glennie, E. J. (2017). Performance-based assessment in the secondary STEM classroom. *Technology and Engineering Teacher, 76*(6), 18–22. https://eric.ed.gov/?id=EJ1133072

Radloff, J., & Guzey, S. (2017). Investigating changes in preservice teachers' conceptions of STEM education following video analysis and reflection. *School Science and Mathematics, 117* 158-167. https://doi.org/10.1111/ssm.12218

Sarıcan, G. (2017). *The effects of integrated STEM education on academic achievement, reflective thinking skills for problem-solving, and retention in learning*. [Master's thesis, İstanbul Aydın University and Yıldız Teknik University]. National Thesis Center.

Saxton, E., Burns, R., Holveck, S., Kelley, S., Prince, D., Rigelman, N., & Skinner, E. A. (2014). A common measurement system for K-12 STEM education: Adopting an educational evaluation methodology that elevates theoretical foundations and systems thinking. *Studies in Educational Evaluation, 40*, 18–35. https://doi.org/10.1016/j.stueduc.2013.11.005

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14. https://doi.org/10.3102/0013189X029007004

Sondergeld, S. A. (2014). Closing the gap between STEM teacher classroom assessment expectations and skills. *School Science & Mathematics, 114*(4), 151-153. https://doi.org/10.1111/ssm.12069

Srinivasan, P. (2015). Assessment in STEM SOS. In A. Şahin (Eds.) A *practice-based Model of STEM Teaching STEM Students on the Stage (SOS)* (pp.111-123). Sense Publishers.

Stewart, C. J. & Cash, W. B. (1985). *Interviewing: Principles and practices* (4th ed.). W.C. Brown Publishers.

Şardağ, M., Ecevit, T., Top, G., Kaya, G. & Çakmakcı, G. (2018). Science and engineering applications. In G. Çakmakcı & A. Tekbıyık (Eds.), *Science education by current curriculum and supported by STEM* (pp. 239-264). Nobel Press.

Şekel, S., (2007). *Evaluation of the new science and technology curriculum in the light of teachers' opinions (Gumushane province example).* [Master's thesis, Karadeniz Teknik University]. National Thesis Center.

Şenel Çoruhlu, T., Nas, S. E., & Çepni, S. (2009). Problems facıng scıence and technology teachers usıng alternative assessment tecniques: Trabzon sample. *Van Yüzüncü Yıl University Fac*https://doi.org/10.1016/j.matt.2020.10.010*ulty of Education Journal, 6*(1), 122-141.

Tao, Y. (2020). Forging aspirant undergraduate scientists and engineers into stellar researchers. *Matter,3*(5), 1383–1386.

Tan, A. L., & Leong, W. F. (2014). Mapping curriculum ınnovation in STEM schools to assessment requirements: tensions and dilemmas. *Theory into Practice, 53*, 11–17. https://doi.org/10.1080/00405841.2014.862113

Tekin Poyraz, G. (2018). *Examining the Kayseri province example in STEM education application and the applicability of distance STEM education.* [Master's thesis, Anadolu University]. National Thesis Center.

Tezcan Şirin, G., Kaval Oğuz, E., & Tüysüz, M. (2022). Investigation of the appropriateness of activities in elementary school science textbooks to the science, technology, engineering, and mathematics approach. *Ankara University Journal of Faculty of Educational Sciences, 5*(1), 37-76. https://doi.org/10.30964/auebfd.863341

The Dayton Regional STEM Center. (2017). *STEM education quality framework.* https://p12summit.wordpress.com/wp-content/uploads/2012/04/03_rowleypreisspinnellfranco1.pdf

Vadeboncoeur, J. (1997). Child development and the purpose of education: A historical context for constructivism in teacher education. In V. Richardson (Ed.), *Constructivist teacher education: Building new understandings* (pp. 15-37). Falmer Press.

Ünsal, S. & Kaba, A. (2022). The characteristics of the skill-based questions and their reflections on teachers and students. *Kastamonu Education Journal, 30*(2), 273-282. https://doi.org/10.24106/kefdergi.753717

World Economic Forum (2020). *The World Economic Forum.* https://www3.weforum.org/docs/WEF_TheGlobalCompetitivenessReport2020.pdf

Zengin, N., Kaya, G., & Pektaş, M. (2020). Investigation of measurement and assessment methods used in STEM-based research. *GUJGEF, 40*(2), 329-355. https://dergipark.org.tr/en/download/article-file/993548

GENİŞLETİLMİŞ ÖZET

STEM eğitiminde becerilerin ölçülmesi önemli olsa da 21. yüzyıl becerileri başta olmak üzere çeşitli becerilerin ölçülmesinde sorunlar bulunmaktadır. Dolayısıyla, bu becerilerin ölçülmesi ve değerlendirilmesi için neler yapılacağının üzerinde durulması gerekmektedir (Çepni, 2018). Alan yazını, STEM eğitiminde becerilerin ölçülmesi için rubrik gibi tasarımlar önerse de henüz STEM eğitiminin doğasına uygun nitelikli ölçme araçları sınırlı kalmaktadır (Saxton vd., 2014). Nitelikli bir STEM eğitiminde ölçme değerlendirmenin sahip olması gereken iki özellik vardır (Potter, 2017; Saxton vd., 2014; Sondergeld, 2014; Srinivasan, 2015). Birincisi, STEM eğitiminde ölçme değerlendirme öğrencilerin bilgi ve becerilerini göstermelerine imkân sağlamalıdır. İkincisi, STEM eğitiminde ölçme değerlendirme çeşitli becerileri ölçebilmelidir. Dolayısıyla, STEM eğitiminde ölçme değerlendirme; bilişsel, duyuşsal, psiko-motor ve 21. yüzyıl becerileri gibi üst düzey düşünme becerilerini belirleyebilmelidir (Harwell vd., 2015; Potter, 2017; Saxton vd., 2014; Sondergeld, 2014; Srinivasan, 2015; Tan ve Leong, 2014; The Dayton Regional STEM Center, 2017). Bu çalışmada sınıflarında STEM etkinliklerini uygulayan fen bilimleri öğretmenlerinin görüşleri alınarak STEM eğitiminde becerileri ölçerken yaşadıkları sorunlarının tespit edilmesi amaçlanmıştır.

Çalışmada, nitel araştırma yönteminin özelliklerine uygun olarak, fenomenoloji deseni kullanılmıştır. Bu çalışmada örneklem çeşitlemesi yapılarak amaçlı örnekleme türlerinden olan ölçüt ve maksimum çeşitlilik örneklemesi kullanılmıştır. Ölçüt olarak fen bilimleri dersinde STEM etkinliklerini uygulayan fen bilimleri öğretmenleri çalışma grubuna seçilmiştir. Çalışma, maksimum çeşitlilik örneklemesi ile cinsiyet ve okul bağlamında STEM uygulamalarında herhangi bir ortak olgunun olup olmadığını ortaya çıkarmaya çalışmıştır. Dolayısıyla, çalışma grubu, 2021-2022 eğitim-öğretim yılında K*** ve N*** illerinde MEB'e bağlı çeşitli resmi, özel ortaokul, BİLSEM ve resmi lisede görev yapmakta olan iki kadın ve beş erkek yedi fen bilimleri öğretmeni seçilerek oluşturulmuştur. Çalışma grubu seçilirken fen bilimleri öğretmenlerinin gönüllü olması ve farklı sınıf seviyelerinde fen bilimleri dersini yürütmeleri dikkate alınmıştır ve gerekli izinler alınmıştır. Görüşme sürecine başlamadan önce alan yazını taraması yapılarak görüşme soruları havuzu oluşturulmuştur. Çalışmanın ikinci yazarı olan fen eğitimcisinden uzman görüşü alınmıştır. Ardından, fen bilimleri öğretmenlerinin STEM etkinliklerini uygulamaları esnasında beceri ölçerken karşılaştıkları sorunları belirlemeye yönelik sorular oluşturulmuştur. Bu kapsamda görüşme formu hazırlanmıştır. Görüşme formu, genel bilgilendirme-etik açıklamalar, kişisel bilgiler ve açık uçlu soruların yer aldığı kısımlardan oluşmaktadır. Bu katılımcılarla iki açık uçlu sorudan oluşan yarı yapılandırılmış görüşmeler gerçekleştirilmiştir. Görüşmeler birinci araştırmacı tarafından çevrimiçi olarak gerçekleştirilmiş ve katılımcılardan izin alınarak ses kaydı alınmıştır. Ses kayıtları yazıya döküldükten sonra kategoriler, alt kategoriler ve kodlar oluşturularak içerik analizi yapılmıştır. Çalışmanın geçerlik ve güvenirlik kontrolleri yapılmıştır.

Çalışmada doğrudan alıntılar yapılarak bulgular sunulmuştur. Yüksek, orta ve düşük frekansa sahip kodlardan birer alıntı verilmiştir. Bulgular, STEM eğitiminde beceri ölçmede yaşanan problemler ve ölçülemeyen beceriler ve gerekçesi kategorileri olmak üzere iki başlık altında incelenmiştir.. STEM eğitiminde beceri ölçmede yaşanan problemler kategorisi kazanım, süre, öğrenci, aktif katılım ve matematik, teknoloji ve mühendislik ile entegrasyon alt kategorilerinden oluşmaktadır. Ölçülemeyen beceriler ve gerekçesi kategorisi tek başlık altında verilmiştir.

Çalışma sonucunda STEM etkinliklerinde beceri ölçerken fen bilimleri öğretmenlerinin kazanım, süre, öğrenci, aktif katılım ve entegrasyon açılarından çeşitli problemlere sahip olduğu ortaya çıkmıştır. Bu problemler arasında, alternatif ölçme aracı bulamama, öğrencilerin beceri sorularını yetiştirememesi, öğrencilerin geleneksel ölçme araçlarına alışkın olması, düşük hazır bulunuşluk, entegrasyonda beceri ölçümü yapılamaması ve sınıfların kalabalık olması vardır. Öğretmenlerin çeşitli becerileri ölçememe gerekçeleri olarak bilişsel becerilere ağırlık verilmesi, yeterli alternatif ölçme aracının olmaması ve ölçeklerin geçerlik ve güvenirlik açısından yetersiz olmaları sonucu ortaya çıkmıştır. Fen bilimleri öğretmenlerine ve öğretmen adaylarına çeşitli alternatif ölçme araçlarını hazırlama ve uygulama eğitimleri verilmesi ve fen bilimleri öğretmenleri STEM etkinliklerinde öğrencilerinin üst düzey düşünme becerilerini geliştirmeye yönelik çalışmalara ağırlık verilmesi önerilmiştir.