Journal of Turkish Chemical Society Section C: Chemistry Education (JOTCSC)
Vol. 8, Issue 2, September 2023, pp. 137-156. ISSN: 2459-1734
Türkiye Kimya Derneği Dergisi Kısım C: Kimya Eğitimi
Cilt 8, Sayı 2, Eylül 2023, sayfa 137-156. ISSN: 2459-1734

Araştırma Makalesi / Research Article

Investigation of the Chemistry and Science Teachers' Self-Efficacy on Alternative Assessment and Evaluation in Terms of Some Variables[1]

Mehtap ÖZDEMİR[1], Canan NAKİBOĞLU[2]

[1] 75th Year Secondary School, Edremit, Balikesir/Türkiye, mehtapozdemir000@gmail.com, https://orcid.org/0000-0002-2102-9457

[2] Balıkesir University, Necatibey Education Faculty, Balıkesir/Türkiye, canan@balikesir.edu.tr, https://orcid.org/0000-0002-7292-9690

Received: 02.09.2023 Accepted: 21.09.2023

Doi: https://doi.org/10.37995/jotcsc.1354350.

Abstract: When the explanations of the 2018 Secondary Education Chemistry and Science Curricula concerning the assessment and evaluation approach are examined, it is seen that the emphasis is on acting with the understanding of maximum diversity and flexibility in the assessment and evaluation process in both curricula. It is stated that in ensuring the effectiveness of assessment and evaluation practices, the priority is on teachers and educational practitioners. It can be said that the assessment and evaluation approach of both the 2018 Secondary Education Chemistry and Science Curricula are largely based on the alternative assessment and evaluation approach. This study investigates the level of self-efficacy of chemistry and science teachers towards alternative assessment and evaluation tools in terms of different variables. A total of 142 teachers, 97 female and 45 male, participated in the study. 32 of these teachers are chemistry teachers and 110 of them are science teachers. At the end of the study, it was determined that the teachers' self-efficacy levels for alternative assessment and evaluation tools were high. It has been revealed that the variables of gender, branch, and years of service do not have a significant effect on teachers' self-efficacy. In addition, it was concluded that the self-efficacy of teachers with doctoral education is higher than that of teachers with undergraduate and graduate degrees. Suggestions are provided at the end of the study.

Keywords: Self-Efficacy, chemistry and science teachers, alternative assessment and evaluation.

1Corresponding author: Mehtap ÖZDEMİR (This study was produced from Mehtap Özdemir's master's thesis and some part of the study was presented in UKEK2023.)

INTRODUCTION

Assessment and evaluation are two indispensable elements of the teaching process. Assessment can be defined as the process of determining the numbers of events, facts, objects, entities and their qualities (variables) that exist in the universe and are the subject of research (Turgut, 1984). Evaluation “is a multi-step, systematic process involving the collection and interpretation of educational data to determine the effectiveness of teaching and learning” (MoNE, 2004). İşman (2006) defines evaluation as expressions that show how well the determined goals have been achieved or not, and if so, to what extent. Evaluation not only gives grades to students but also allows students to see the processes they go through and their shortcomings (MoNE, 2018a and 2018b).

The change that has taken place in the curriculum of science courses in Türkiye in recent years has also revealed the necessity of using strategies in which the student is more active and in the center and organizing appropriate environments for these strategies. Thus, programs have adopted a constructivist approach. The constructivist understanding takes into account the individual differences of students and suggests that students construct new information in a unique way by adding it to the information they already have (Bodner, 1986). According to the constructivist approach, students should be offered multiple assessment and evaluation opportunities. On the other hand, it is seen that Turkish teachers mostly focus on traditional measurement tools for assessment and evaluation purposes at both high school and secondary school levels (Bayat & Şentürk, 2015; Özenç, 2013; Sütçü & Bulut, 2017). Şenel et al. (2018) conducted a study to reveal the issues that chemistry teachers feel inadequate in the field of measurement and evaluation and the problems they experience in this field. As a result of this study, it was determined that approximately one quarter of chemistry teachers felt inadequate in applying alternative measurement and evaluation methods.

It is clear that the effectiveness of alternative assessment and evaluation techniques, which are frequently used in the program, depends on the individual teachers (Cheng, 2006). Karaaslan (2015) stated that the reason why science and technology teachers and classroom teachers mostly use traditional assessment and evaluation tools is that the teachers do not know enough about alternative assessment and evaluation techniques, it takes more time, and the classrooms are crowded. Cheng (2006), in his study with science teachers in Hong Kong secondary schools on the use of alternative assessment tasks in their lessons, also questioned the reason for the problems in teachers' use of them. He gathered the teachers' opinions about the difficulties and anxieties in implementing alternative assessment tasks under six themes. These are as follows: lack of experience or knowledge, teaching time, teaching load, developing consensus among teachers, students' skills and abilities, and low priority for students. From this explanation, it can be said that the teachers' lack of knowledge about these alternative assessment and evaluation techniques is directly connected to their pedagogical content

knowledge (PCK) which was introduced by Shulman (1986) being one of the content knowledge that a teacher should have. In the model offered by Magnusson et al. (1999), PCK consists of five components. Teachers' assessment knowledge is one of these five types of knowledge. It refers to the knowledge that the teacher has about the measurement that he performs to get feedback from the students during or at the end of the teaching. The self-efficacy of teachers to adopt different assessment and evaluation techniques or to use different assessment approaches is also related.

Bandura (1995) stated that there are four main sources of self-efficacy beliefs. While the most effective one among them is shown as the information gained by individuals directly from their own experiences, other sources are individuals' observations of successful or unsuccessful practices, the influence of society on success, and the psychological state of success (cited in Üztemur & Metin, 2015). Bandura (1982) also expressed that one's competence in dealing with one's environment is not a matter of fixed action or just knowing what to do, but that cognitive, social, and behavioral skills involve a productive ability that must be organized in integrated courses of action to serve innumerable purposes. Accordingly, he stated that perceived self-efficacy is related to judgments about how well a person can apply the necessary action plans to cope with possible situations. In addition, the perception of self-efficacy is effective on emotional reactions as well as behaviors. Self-efficacy is defined as the idea of an individual's ability to control his own behavior, his beliefs about whether he will be able to overcome the event, the difficulties he will encounter, and his judgment about himself (Semenoğlu, 2015:234; cited in Kansu & Sayar, 2018).

Since these characteristics of people with high self-efficacy cause a positive impact in every aspect of their lives, a relationship can be established between teachers' self-efficacy and teaching practices. Many researchers state that the quality of assessment and evaluation practices in education largely depends on teachers' self-efficacy in this field (Baş & Beyhan, 2016; Çakan, 2004; Kılınç, 2011). Individuals with high self-efficacy are also quite competent in using alternative measurement and evaluation tools (Yayla, 2011). As pointed out by Atılgan, Kan, and Doğan (2007), in order to reveal the effectiveness of the applied teaching programs, measuring the competency that teachers should have in order to evaluate students at the beginning, during the process, and at the end of the process, to determine to what extent they have achieved the determined goals and to reveal the deficiencies and evaluation of self-efficacy (Cited in Baş & Beyhan, 2016). When the explanations included in the assessment and evaluation approach of both the Turkish 2018 Secondary Education Chemistry Curriculum and Science Curriculum and the items of the principles that guide the assessment and evaluation techniques in the curricula are examined, it is seen that the assessment and evaluation approach of the programs largely requires alternative measurement and evaluation. For this reason, it is important to determine the self-efficacy of chemistry and science

teachers towards assessment and evaluation and the factors related to their self-efficacy. Şahin and Kaya (2020) analyzed the content of theses and articles about alternative assessment and evaluation published in YÖK National Thesis Center and DergiPark according to the determined criteria. 27 theses and 18 articles between 2008 and 2020 were selected by purposeful sampling method were subjected to content analysis. As a result of the analysis, it was determined that 4 theses and 5 articles from these studies were related to the self-efficacy of teachers or teacher candidates toward alternative measurement and evaluation. Although there are a limited number of studies examining the self-efficacy of science teachers for alternative assessment and evaluation tools (Tatar & Buldur, 2013), the literature review showed that no study has been conducted that examines both the self-efficacy of chemistry teachers and the self-efficacy of chemistry and science teachers together. Therefore, in this study, the self-efficacy of chemistry and science teachers was examined in terms of the effects of different variables in line with the following research questions.

1. Do the self-efficacy of Chemistry and Science teachers, regarding alternative assessment and evaluation tools, show a significant difference according to gender?
2. Do the self-efficacy of Chemistry and Science teachers, regarding alternative assessment and evaluation tools, show a significant difference according to their branch (science or chemistry)?
3. Do the self-efficacy of Chemistry and Science teachers, regarding alternative assessment and evaluation tools, show a significant difference according to teachers' years of service?
4. Do the self-efficacy of Chemistry and Science teachers, regarding alternative assessment and evaluation tools, show a significant difference according to the type of faculty graduated?
5. Do the self-efficacy of Chemistry and Science teachers, regarding alternative assessment and evaluation tools, show a significant difference according to the graduation level of teachers?
6. Do the self-efficacy of Chemistry and Science teachers, regarding alternative assessment and evaluation tools, show a significant difference according to taking lessons for alternative assessment and evaluation during the undergraduate education of teachers?

METHOD

The study's model of the study, sample, and the path followed in data gathering and analysis of data are explained below.

Model of the Study

In this study, the *general survey model*, one of the survey models, was used. Survey models are research approaches that aim to describe a past or present situation as it exists. The event, individual or object that is the subject of the research is tried to be defined in its own conditions and as it exists. No effort has been made to change or influence them in any way (Karasar, 2011, p.77).

Sample

In the study, sampling was carried out according to *convenient sampling*, which is one of the purposive sampling methods. In convenient sampling, the researcher starts to create his sample starting from the most accessible respondents until he reaches a group of the size he/she needs, or s/he works on a situation or example that is the most accessible and will provide maximum savings (Büyüköztürk et al., 2017). Before starting the study, necessary ethical approval was obtained from the Science and Engineering Sciences Research Ethics Committee of Balikesir University. During the data collection phase, consent was requested from the teachers for voluntary participation and data were collected from the teachers who gave consent. At the stage of data collection, a pilot study was conducted and the data were collected both through Google forms and by reaching the teachers face to face. The sample of the research consists of 142 chemistry and science teachers in high schools and secondary schools in different provinces and districts in Türkiye and demographic information of chemistry and science teachers is shown in Table 1.

Data Gathering

Data collection in this study was done using an *Alternative Assessment and Evaluation Self-Efficacy Scale* (AAESE Scale). In the literature, there are scales developed for different teacher groups regarding the self-efficacy of alternative assessment and evaluation tools. As a result of examining these scales, it was decided to use the scale developed by Aksoy (2018). After obtaining the necessary permission, the reliability of the AAESE Scale was tested with a pilot study. A total of 61 teachers, including 39 female and 22 male, 16 chemistry and 45 science teachers, participated in the pilot study. As a result of the pilot study, the Cronbach alpha reliability coefficient of the AAESE Scale was determined as 0.96. After it was decided that this reliability coefficient was appropriate, the application for the AAESE Scale was carried out.

The AAESE Scale, which includes 20 items in 5-point Likert type, consists of two factors. The first factor was named as "*preparation and use of alternative measurement tools*" and the second factor as "*problems encountered with alternative measurement tools and providing solutions*" by the developer. These two factors are generally defined as the level of self-efficacy for the whole scale. The application was made by adding the

first part in which demographic information concerning teachers was questioned in front of this part, which included AAESE Scale questions.

Table 1

Demographic information of chemistry and science teachers.

Variables		f	%
Gender	Female	97	68.0
	Male	45	32.0
Type of school served	Secondary school	107	75.4
	High school	35	24.6
Branch	Science	110	77.5
	Chemistry	32	22.5
Years of service	Less than 1 year	2	1.4
	1-5 years	12	8.5
	6-10 years	28	19.7
	11-20 years	68	47.9
	Over 20 years	32	22.5
Location of the school	Province	26	18.3
	District	106	74.6
	Village	9	6.3
	Province and district	1	0.7
Type of school studied	State school	122	85.9
	Private school	18	12.7
	Private course centre	2	1.4
Type of faculty graduated	Faculty of Education	123	86.6
	Faculty of Arts and Sciences	18	12.7
	Engineering	1	0.7
Graduation level	Bachelor's degree	99	69.7
	Ms	37	26.1
	PhD	6	4.2
Taking courses concerning alternative assessment and evaluation	Yes	75	52.8
	No	67	47.2

Analysis of the Data

In the analysis of the data, first of all, descriptive statistics regarding the responses of the teachers to the AAESE were calculated. Then, according to some variables, it was investigated whether there was a significant difference between teachers' self-efficacy.

In order to investigate the effect of independent variables on the self-efficacy level of the whole scale, factor 1 and factor 2, it was examined whether the data obtained from the AAESE showed normal distribution. Firstly, skewness and kurtosis values were examined. It was also checked with the Kolmogorov Smirnov Z test. Skewness/standard error values for factor 1 and factor 2 with the whole scale were calculated as 4.05, 2.74 and 4.82, respectively. The kurtosis/standard error values were found to be 7.93, 5.05 and 7.89, respectively. It was determined that the scale data did not show normal distribution since these values were not in the range of +1.96 to -1.96 points. According to the result of the

Kolmogorov-Smirnov test, it is understood that the p<.005 (p=.000) scale data do not show a normal distribution. Therefore, in the next step, Mann-Whitney U and Kruskal-Wallis H tests, which are non-parametric tests, were used in the analysis of the data in order to answer the sub-problems of the study related to self-efficacy.

FINDINGS

The findings obtained from the study are given below to answer the research questions.

Findings Regarding the First Research Question

Table 2 presents the descriptive statistics results, which were conducted to answer the first research problem in which chemistry and science teachers' self-efficacy levels for alternative assessment and evaluation were investigated.

Table 2

Descriptive statistics results of teachers' self-efficacy towards alternative assessment and evaluation tools.

	Total Self-efficacy	Self-efficacy for Factor 1*	Self-efficacy for Factor 1**
N	142	142	142
Scale items	1-20	1-8, 10, 14	9, 11-13, 15-20
Mean	70.04	34.27	35.77
Standard deviation	1.35	7.40	6.74
Variance	183.16	54.72	45.46
Minimum	20	10	10
Maximum	100	50	50
Cronbach α	.96	.94	.92

*Preparation and use of alternative assessment tools.
**Problems encountered with alternative assessment tools and finding solutions.

As seen in Table 2, the Cronbach α for the total self-efficacy level was calculated as .96, and the mean was found as 70.04. Since the highest value that can be taken from the whole scale is 100, these mean show that teachers' total self-efficacy level towards alternative assessment and evaluation tools is high. The Cronbach α was calculated as .94 and the mean as 34.27 for factor 1, named *preparation and use of alternative assessment tools.* For factor 2, titled *problems encountered with alternative assessment tools and finding solutions,* the Cronbach's α was calculated as .92 and the mean was found as 35.77. Since the highest value that can be taken for each factor is 50, these means show that teachers' self-efficacy level is high for both factors.

Findings Regarding the Second Research Question

In the second research question, whether gender has an effect on teachers' self-efficacy was investigated for both the scale and both factors of the scale. The data obtained from the self-efficacy according to the gender variable are presented in the mean rank and sum of ranks Table 3, and the Mann-Whitney U test results are presented in Table 4.

Table 3

Ranks obtained from the self-efficacy scale according to the gender variable.

	Gender	N	Mean Rank	Sum of Ranks
Scale	Female	97	72.94	7075.50
	Male	45	68.39	3077.50
Factor 1 of scale	Female	97	71.46	6932.00
	Male	45	71.58	3221.00
Factor 2 of scale	Female	97	73.86	7164.00
	Male	45	66.42	2989.00

Table 4

Test Statistics on gender variable.

	Scale	Factor 1 of scale	Factor 2 of scale
Mann-Whitney U	2.042E3	2.179E3	1.954E3
Wilcoxon W	3.078E3	6.932E3	2.989E3
Z	-.614	-.015	-1.004
Asymp. Sig. (2-tailed)	.539	.988	.315

From Table 3, it is seen that the total self-efficacy scores of female teachers for the whole scale are higher than those of male teachers. In addition, the self-efficacy scores of female and male teachers related to the first factor "preparation and use of alternative measurement tools" are almost equal. The self-efficacy scores of female teachers related to the second factor "problems encountered with alternative measurement tools and finding solutions" are higher than those of male teachers. As a result of the Mann-Whitney U test (Table 4), which was performed to determine whether there was a significant difference between these scores, it was found that $p>.05$ ($p_{scale}=.539$; $p_{F1}=.988$; $p_{F2}=.315$) for the scale, factor 1 and factor 2, respectively. These p values show that there is no significant difference between the total self-efficacy scores of female and male teachers for the whole scale, and between female and male teachers' self-efficacy scores for both factors.

Findings Regarding the Third Research Question

In the third research question, whether the branch (science or chemistry) has an effect on teachers' self-efficacy was investigated for both the scale and both factors of scale. The mean Rank and Sum of Ranks of the data obtained from the self-efficacy scale according to the branch variable are presented in Table 5, and the Mann-Whitney U test results are presented in Table 6.

Table 5

Ranks obtained from the self-efficacy scale according to the branch variable.

	Branch	N	Mean Rank	Sum of Ranks
Scale	Science	110	68.49	7534.00
	Chemistry	32	81.84	2619.00
Factor 1 of scale	Science	110	67.67	7444.00
	Chemistry	32	84.66	2709.00
Factor 2 of scale	Science	110	70.33	7736.50
	Chemistry	32	75.52	2416.50

Table 6

Test Statistics on branch variable.

	Scale	Factor 1 of scale	Factor 2 of scale
Mann-Whitney U	1.429E3	1.339E3	1.632E3
Wilcoxon W	7.534E3	7.444E3	7.736E3
Z	-1.617	-2.059	-.629
Asymp. Sig. (2-tailed)	.106	.040	.530

In Table 5, it is seen that the total self-efficacy scores of chemistry teachers for the scale are higher than that of science teachers. Similarly, the self-efficacy scores of chemistry teachers regarding the factor 1, "preparation and use of alternative measurement tools", and the factor 2, "problems encountered with alternative measurement tools and finding solutions", are higher than those of science teachers. As a result of the Mann-Whitney U test (Table 6) performed to determine whether there was a significant difference between these scores, $p>.05$ ($p_{scale}=.106$; $p_{F2}=.530$) for the scale and the factor 2, respectively. These p values show that there is no significant difference between the self-efficacy scores of chemistry and science teachers for the scale and the 2nd factor of the scale. It was found to be $p<.05$ ($pF_1=.040$) for the factor 1 of the scale, and this value shows that the branch has an effect in favor of chemistry teachers on teachers' self-efficacy for the 1st factor.

Findings Regarding the Fourth Research Question

In the fourth research question, the effect of years of service on teachers' self-efficacy was investigated for both the scale and both factors of it. The mean rank and sum of ranks of the data obtained from the self-efficacy scale according to the service year variable are presented in Table 7, and the Kruskal-Wallis H Test results are presented in Table 8.

Table 7

Ranks obtained from the self-efficacy scale according to the years of service *variable.*

	Years of service	N	Mean Rank	Sum of Ranks
Scale	Less than 1 year	2	68.50	2
	1-5 years	12	78.17	12
	6-10 years	28	68.91	28
	11-20 years	68	71.82	68
	Over 20 years	32	70.78	32
Factor 1 of scale	Less than 1 year	2	67.75	2
	1-5 years	12	73.21	12
	6-10 years	28	71.89	28
	11-20 years	68	71.31	68
	Over 20 years	32	71.16	32
Factor 2 of scale	Less than 1 year	2	66.25	2
	1-5 years	12	81.92	12
	6-10 years	28	67.55	28
	11-20 years	68	72.89	68
	Over 20 years	32	68.42	32

Table 8

Test Statistics on years of service *variable.*

	Scale	Factor 1 of scale	Factor 2 of scale
Chi-Square	.451	.044	1.322
df	4	4	4
Asymp. Sig.	.978	1.000	.858

Table 7 shows that the total self-efficacy scores of teachers with 1-5 years of professional service for the scale, factor 1 and factor 2 are higher than those of teachers working in other years of professional service. As a result of the Kruskal-Wallis H Test (Table 8), $p>.05$ was found for the scale, factor 1 and factor 2 of the scale ($p_{scale}=.978$; $p_{F1}=1.00$; $p_{F2}=.858$). This finding reveals that years of professional service have no effect on

teachers' self-efficacy for alternative assessment and evaluation tools for the whole scale and of both factors.

Findings Regarding the Fifth Research Question

In the fifth research question, the effect of the type of faculty graduated on teachers' self-efficacy was investigated for both the scale and both factors of the scale. The mean rank and sum of the ranks of the data obtained from the self-efficacy scale according to the type of faculty graduated variable are presented in Table 9, and the Kruskal-Wallis H Test results are presented in Table 10.

Table 9

Ranks obtained from the self-efficacy scale according to the type of faculty graduated *variable.*

	Type of faculty graduated	N	Mean Rank	Sum of Ranks
Scale	Faculty of Education	123	68.04	123
	Faculty of Arts and Sciences	18	94.22	18
	Engineering Faculty	1	88.50	1
Factor 1 of scale	Faculty of Education	123	69.15	123
	Faculty of Arts and Sciences	18	84.53	18
	Engineering Faculty	1	125.50	1
Factor 2 of scale	Faculty of Education	123	68.08	123
	Faculty of Arts and Sciences	18	92.81	18
	Engineering Faculty	1	108.50	1

Table 10

Test Statistics on the type of faculty graduated variable.

	Scale	Factor 1 of scale	Factor 2 of scale
Chi-Square	6.556	3.945	6.496
df	2	2	2
Asymp. Sig.	.038	.139	.039

Table 9 shows that the total self-efficacy scores of teachers who graduated from the Faculty of Arts and Sciences for the scale are higher than those of the teachers who graduated from other faculties. For factor 1 and factor 2, the self-efficacy scores of a teacher who graduated from the Faculty of Engineering are higher than those of the teachers who graduated from other faculties. As a result of the Kruskal-Wallis H Test (Table 10), $p<.05$ for the scale and the factor 2 was found ($p_{scale}=.038$; $p_{F2}=.039$). This

shows the type of faculty graduated has an effect on teachers' self-efficacy for alternative assessment and evaluation tools for scale and factor 2. While this significant difference in self-efficacy for the scale is in favor of the teachers who graduated from the Faculty of Arts and Sciences, this significant difference regarding the problems encountered with the alternative measurement tools, which is the second factor, and the self-efficacy for providing solutions, is in favor of the teachers who graduated from the Faculty of Engineering.

Findings Regarding the Sixth Research Question

In the sixth research question, the effect of the graduation level of teachers on their self-efficacy was investigated for both the scale and both factors of it. The mean rank and sum of the ranks of the data obtained from the self-efficacy scale according to the graduation level of teachers' variable are presented in Table 11, and the Kruskal-Wallis H Test results are presented in Table 12.

Table 11

Ranks obtained from the self-efficacy scale according to the graduation level of teachers.

	The graduation level	N	Mean Rank	Sum of Ranks
Scale	Bachelor's degree	99	64.36	99
	Ms	37	81.47	37
	PhD	5	124.90	5
Factor 1 of scale	Bachelor's degree	99	66.45	99
	Ms	37	78.36	37
	PhD	5	106.50	5
Factor 2 of scale	Bachelor's degree	99	65.13	99
	Ms	37	80.15	37
	PhD	5	119.50	5

Table 12

Test Statistics on the graduation level of teachers variable

	Scale	Factor 1 of scale	Factor 2 of scale
Chi-Square	13.797	6.231	10.964
df	2	2	2
Asymp. Sig.	.001	.044	.004

Table 11 shows that the total self-efficacy scores of teachers having Ph.D. degrees for the scale, factor 1 and factor 2 are higher than those of teachers having Bachelor's degrees and Ms Degrees. As a result of the Kruskal-Wallis H Test (Table 12), it was found to be

$p<.05$ for the scale and both factors ($p_{scale}=.001$; $p_{F1}=.044$; $p_{F2}=.004$). This shows that the level of education has an effect on teachers' self-efficacy towards alternative measurement and evaluation tools for the scale and both factors of the scale. This significant difference regarding self-efficacy for the entire scale and both factors is in favor of teachers with doctoral education level.

Findings Regarding the Seventh Research Question

In the seventh research question, the effect of taking courses for alternative assessment and evaluation during the undergraduate education of teachers on their self-efficacy was investigated for both the scale and the first and second factors. The mean rank and sum of the ranks of the data obtained from the self-efficacy scale according to the taking courses for alternative assessment and evaluation during the undergraduate education of the teachers are presented in Table 13 and the Mann-Whitney U Test results are presented in Table 14.

Table 13

Ranks obtained from the self-efficacy scale according to the courses taken variable.

	The courses taken	N	Mean Rank	Sum of Ranks
Scale	Yes	75	73.93	5544.50
	No	67	68.78	4608.50
Factor 1 of scale	Yes	75	75.65	5674.00
	No	67	66.85	4479.00
Factor 2 of scale	Yes	75	72.57	5442.50
	No	67	70.31	4710.50

Table 14

Test Statistics on the courses taken variable.

	Scale	Factor 1 of scale	Factor 2 of scale
Mann-Whitney U	2330E3	2.201E3	2.432E3
Wilcoxon W	4.608E3	4.479E3	4.710E3
Z	-.744	-1.275	-.328
Asymp. Sig. (2-tailed)	.457	.202	.743

Table 13 shows that the total self-efficacy scores of teachers who took courses for alternative assessment and evaluation during undergraduate education for the scale, factor 1 and factor 2 are higher than those of teachers who did not take these types of courses. As a result of the Mann-Whitney U test (Table 14), $p>.05$ for the scale and for

both factor 1 and 2 (p_{scale}=.457; p_{F1}=.202; p_{F2}=.743). These p values show that there is no significant difference between the self-efficacy of the teachers who took and did not take courses at the undergraduate level for the scale and the first and second factors.

RESULTS AND DISCUSSION

At the end of this study, in which the self-efficacy of Chemistry and Science teachers towards alternative measurement and evaluation tools was examined, it was determined that teachers' self-efficacy towards alternative measurement and evaluation tools was high. The first dimension of the scale, whose reliability level is calculated to be quite high, is "Preparation and use of alternative measurement tools". When the teachers' self-efficacy score averages for this dimension were examined, it was determined that the level of self-efficacy was high. It was also determined that the level of self-efficacy was high for the second dimension of the scale, “Problems encountered with alternative measurement tools and finding solutions”. It was determined that teachers' self-efficacy for the 2nd dimension was higher than the 1st dimension.

Although it was found that the total self-efficacy scores of female teachers for the whole scale and the second factor of the scale were higher than those of male teachers; it was concluded that there was no significant difference between the self-efficacy scores of female and male teachers.

It was found that the total self-efficacy scores of chemistry teachers for the scale and both factors of the scale were higher than those of science teachers' total self-efficacy scores. On the other hand, while it was determined that this difference was not significant for scale and factor 2, it was found to be significant for the first factor of the scale, "preparation and use of alternative measurement tools". From this point of view, it was seen that chemistry teachers' preparation and use of alternative measurement tools were higher than that of science teachers. As a result, there are different subjects such as physics, chemistry, and biology among the teaching subjects of science teachers. The fact that one of these fields is included more in the courses they take for measurement and evaluation may cause them to have deficiencies in preparing alternative measurement tools for all science subjects. Many studies in the literature have revealed that teachers' knowledge levels have a significant impact on their preparation and use of alternative assessment tools (Cheng, 2006; Flowers et al., 2005; Karaaslan, 2015).

Although the difference between the teachers in their professional years of service is not significant, it was concluded that the teachers with the highest self-efficacy level for the total self-efficacy level and for both factors were the teachers whose years of service were between 1-5 years. A similar result was reached in the study conducted by Watt (2005) with mathematics teachers in Australia. Watt (2005) determined that teachers generally do not use alternative assessment and evaluation tools but he found that the

situation was different for teachers with less professional experience and they use alternative assessment and evaluation tools. This situation can also be associated with teachers' knowledge of alternative assessment and evaluation tools. It can be said that in the first years of the profession, teachers' undergraduate knowledge is fresh and they can better remember a lot of information they have learned in their courses.

It was concluded that the total self-efficacy scores of teachers who graduated from the Faculty of Arts and Sciences for the scale were higher than those of the teachers who graduated from other faculties and that this difference was significant. Another important result reached in the study is the total self-efficacy scores of teachers having Ph.D. degrees for the scale, factor 1 and factor 2 are higher than those of teachers having Bachelor's Degrees and Ms Degrees and this difference is significant.

Another result reached in the study is related to the total self-efficacy scores of teachers who took courses for alternative assessment and evaluation during undergraduate education. Although it was determined that the teachers who took courses in undergraduate education were found to have high self-efficacy for both scale and factors, it was concluded that there was no significant difference between the self-efficacy of the teachers who took and did not take courses at the undergraduate level for the scale and both factors.

The results obtained from the study and studies conducted with teachers from different branches in the literature (Bayat & Şentürk, 2015; Cheng, 2006; Flowers et al., 2005; Şenel Çoruhlu et al., 2009; Şenel et al., 2018) show that teachers' use of alternative assessment and evaluation tools and strategies in the classroom is related to their field and field education knowledge. For this reason, the most important recommendation is related to the undergraduate teaching programs of teacher candidates. In the study conducted by Tatar and Buldur (2013) with pre-service science teachers, the self-efficacy of pre-service teachers was increased by the teaching program, which also supports this idea. Although Nakiboğlu and Karakoç (2005) have dealt with three types of knowledge in the studies with the knowledge that a teacher should have for many years in Türkiye, the types of knowledge are "content knowledge", "teaching profession knowledge" and "general cultural knowledge" stated that “pedagogical content knowledge (PCK)”, which is the fourth knowledge that a teacher should have, has been added to these three types of knowledge in recent years. The importance of the PCK is understood and the courses aimed at gaining this type of knowledge are added to teacher training programs recently in Türkiye. However, it was seen that "chemistry content courses" and "chemistry education courses” which is related to PCK are coded as "chemistry courses" under the same group in the fixed teacher training programs prepared by YÖK in 2018. This is still like three types of knowledge in the training of teachers in Türkiye. It can be said that it supports an idea. In addition, in teacher training programs, "chemistry content courses" and "chemistry education" courses are considered only as the "content knowledge" of the

teacher, and "PCK" seems not to be taken into account, and considering these two groups together considerably reduces the time allocated to these courses. For this reason, it is necessary to realize that chemistry education knowledge is "PCK" and that its teaching is different from chemistry content knowledge. Accordingly, PCK course hours of chemistry and science teachers in undergraduate teaching programs should be increased. In addition, in the study, the high self-efficacy of chemistry teachers graduated from the Faculty of Science and Literature can be associated with the chemistry knowledge they received in their education. This situation again reveals that the hours of chemistry content courses, which were reduced in teaching undergraduate programs, should also be increased.

Another suggestion can be made for teachers in the profession. In-service training courses should be organized for Chemistry and Science teachers on what alternative assessment and evaluation are in their fields, which measurement tools are available for this type of assessment and evaluation, and how they can be prepared and used.

A final suggestion is that the updates made in secondary and primary education curricula can be harmonized with national examinations, as national examinations have a great influence on which methods and materials teachers will use in their classroom practices in Türkiye.

Conflict of Interest Declaration

The authors have not declared a potential conflict of interest during the research, authorship, and publishing of this article.

Support / Financing Information

The authors have not received any financial support during research, authorship, and publishing of this article.

Ethical Committee Decision / Permission

Ethical approval for this research was received at the meeting of Balıkesir University Science and Engineering Sciences Scientific Research Ethics Commission dated 17.03.2022 and numbered 2022/2.

Kimya ve Fen Bilimleri Öğretmenlerinin Alternatif Ölçme ve Değerlendirmeye ilişkin Öz Yeterliliklerinin Bazı Değişkenler Açısından İncelenmesi[2]

Mehtap ÖZDEMİR[1], Canan NAKİBOĞLU[2]

[1] 75. Yıl Orta Okulu, Edremit, Balikesir/Türkiye mehtapozdemir000@gmail.com, https://orcid.org/0000-0002-2102-9457

[2] Balıkesir Üniversitesi, Necatibey Eğitim Fakültesi, Balıkesir/Türkiye, canan@balikesir.edu.tr, https://orcid.org/0000-0002-7292-9690

Gönderme Tarihi: 02.09.2023 Kabul Tarihi: 21.09.2023

Doi: https://doi.org/10.37995/jotcsc.1354350.

Özet: 2018 Ortaöğretim Kimya ve Fen Bilimleri Dersi Öğretim Programlarının ölçme ve değerlendirme yaklaşımına ilişkin açıklamaları incelendiğinde, her iki öğretim programında da ölçme ve değerlendirme sürecinde maksimum çeşitlilik ve esneklik anlayışıyla hareket edilmesine vurgu yapıldığı görülür. Ölçme ve değerlendirme uygulamaları etkililiğinin sağlanmasında önceliğin öğretmenler ve eğitim uygulayıcılarında olduğu da belirtilmektedir. Hem 2018 Ortaöğretim Kimya hem de Fen Bilimleri Dersi Öğretim Programlarının ölçme ve değerlendirme yaklaşımının büyük ölçüde alternatif ölçme ve değerlendirme yaklaşımına dayandığı söylenebilir. Bu çalışma kimya ve fen bilimleri öğretmenlerinin alternatif ölçme ve değerlendirmeye yönelik öz yeterlik düzeylerini farklı değişkenler açısından araştırmaktadır. Araştırmaya 97'si kadın, 45'i erkek olmak üzere toplam 142 öğretmen katılmıştır. Bu öğretmenlerin 32'si kimya öğretmeni, 110'u ise fen bilimleri öğretmenidir. Araştırma sonucunda öğretmenlerin alternatif ölçme ve değerlendirme araçlarına yönelik öz yeterlik düzeylerinin yüksek olduğu tespit edilmiştir. Cinsiyet, branş ve hizmet yılı değişkenlerinin öğretmen öz-yeterliği üzerinde anlamlı bir etkisinin olmadığı ortaya çıkmıştır. Ancak doktora eğitimi almış öğretmenlerin öz yeterliliklerinin lisans ve yüksek lisans mezunu öğretmenlere göre daha yüksek olduğu sonucuna ulaşılmıştır. Çalışmanın sonunda önerilere yer verilmiştir.

Anahtar kelimeler: Öz yeterlilik, Kimya ve Fen Bilimleri öğretmeni, alternatif ölçme ve değerlendirme.

2 Sorumlu yazar: Mehtap ÖZDEMİR (Bu çalışma Mehtap Özdemir'in yüksek lisans tezinden üretilmiş olup, çalışmanın bir kısmı UKEK2023'te sözlü bildiri olarak sunulmuştur.)

REFERENCES

Aksoy, A. (2018). *İlkokul ve ortaokul öğretmenlerinin alternatif ölçme araçları öz yeterliklerinin incelenmesi* [Examination of primary and secondary school teachers' self-efficacy in alternative assessment tools]. Yayımlanmamış yüksek lisans tezi, Süleyman Demirel Üniversitesi, Isparta.

Bandura, A. (1982). *Self-efficacy mechanism in human agency.* American psychologist, 37(2), 122-147. doi:10.1037/0003-066x.37.2.122.

Bandura, A. (1995). Comments on the crusade against the causal efficacy of human thought. *Journal of behavior therapy and experimental psychiatry, 26*(3), 179-190.

Baş, G., & Beyhan, Ö. (2016). Öğretmenlerin eğitimde ölçme ve değerlendirmeye yönelik özyeterlik algılarının bazı değişkenler açısından incelenmesi [Examination of teachers' self-efficacy perceptions on educational measurement and evaluation in terms of some variables]. *Journal of Measurement and Evaluation in Education and Psychology, 7*(1), 18-32.

Bayat, S., & Şentürk Ş. (2015). Fizik, kimya, biyoloji ortaöğretim alan öğretmenlerinin alternatif ölçme değerlendirme tekniklerine ilişkin görüşleri [Physics, Chemistry, Biology Teachers' Views on Alternative Assessment and Evaluation Techniques in Secondary School]. *Amasya Üniversitesi Eğitim Fakültesi Dergisi, 4*(1), 118-135.

Bodner, G. M. (1986). Constructivism: A theory of knowledge. *Journal of chemical education, 63*(10), 873.

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2017). *Bilimsel araştırma yöntemleri* (23. Edition). Pegem Yayıncılık, Ankara.

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

Cheng, M. H. (2006). Junior secondary science teachers' understanding and practice of alternative assessment in Hong Kong: Implications for teacher professional development. *Canadian Journal of Science, Mathematics and Technology Education, 6*(3), 227-243. https://doi.org/10.1080/14926150609556699

Flowers, C., Ahlgrim-Delzell, L., Browder, D., & Spooner, F. (2005). Teachers' perceptions of alternate assessments. *Research and Practice for Persons with Severe Disabilities, 30*(2), 81-92.

İşman, A. (2006). *Öğretimde planlama ve değerlendirme*. Pegem A Yayıncılık, Ankara.

Kansu, A., & Sayar, G. (2018). Özyeterlilik, yaşam anlamı ve yaşam bağlılığı kavramları üzerine bir inceleme. *Üsküdar Üniversitesi İletişim Fakültesi Akademi Dergisi, 1*(1), 1-11.

Karaaslan, O. (2015). *Fen ve teknoloji öğretmenlerinin alternatif ölçme ve değerlendirme tekniklerini uygulamadaki yeterlilikleri*. Yüksek Lisans Tezi. Yüzüncü Yıl Üniversitesi Eğitim Bilimleri Enstitüsü, Van.

Karasar N. (2011). *Bilimsel araştırma yöntemi.* Nobel Yayın Dağıtım, Ankara.

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

Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds*.), Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 95-132). Dordrecht, the Netherlands: Kluwer Academic.

Ministry of National Education (MoNE) (2004). *İlköğretim Fen ve Teknoloji Dersi (4–5. Sınıflar) Öğretim Programı.* Talim ve Terbiye Kurulu Başkanlığı, Ankara.

Ministry of National Education (MoNE) (2018a). *Ortaöğretim Kimya Dersi Öğretim Programı (9, 10, 11 ve 12. Sınıflar).* Talim ve Terbiye Kurulu Başkanlığı, Ankara.

Ministry of National Education (MoNE) (2018b). *Fen Bilimleri Dersi Öğretim Programı (İlkokul ve Ortaokul 3, 4, 5, 6, 7 ve 8. Sınıflar).* Talim ve Terbiye Kurulu Başkanlığı, Ankara.

Nakiboglu, C., & Karakoç, Ö. (2005). The forth knowledge domain a teacher should have: The pedagogical content knowledge. *Kuram ve Uygulamada Eğitim Bilimleri / Educational Sciences: Theory & Practice, 5*(1), 181-206.

Özenç, M. (2013). *Sınıf öğretmenlerinin alternatif ölçme ve değerlendirme yeterliklerinin incelenmesi.* Yayınlanmamış doktora tezi, Marmara Üniversitesi Eğitim Bilimleri Enstitüsü, İstanbul.

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

Sütçü, N. D., & Bulut İ. (2017). Ortaokul matematik öğretmenlerinin alternatif ölçme ve değerlendirme tekniklerine ilişkin yeterlik algıları ve bu teknikleri kullanma düzeyleri. *Dicle Üniversitesi Ziya Gökalp Eğitim Fakültesi Dergisi, 29*, 289-308.

Şahin, Ç., & Kaya, G. (2020). Alternatif ölçme değerlendirme ile ilgili yapilan araştirmalarin incelenmesi: bir içerik analizi. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi, 10*(2), 798-812.

Şenel Çoruhlu, T., Nas, S. E., & Çepni, S. (2009). Fen ve teknoloji öğretmenlerinin alternatif ölçme ve değerlendirme tekniklerini kullanmada karşılaştıkları problemler: Trabzon örneği. [problems facing science and technology teachers using alternative assesment tecnics: Trabzon sample]. *Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi, 6*(1), 122-141.

Şenel, S., Pekdağ, B., & Günaydin, S. (2018). Kimya öğretmenlerinin eğitimde ölçme ve değerlendirme süreçlerinde yaşadıkları problemler ve yetersizlikler. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi, 12*(1), 419-441.

Turgut, M. F. (1984). *Eğitimde Ölçme ve Değerlendirme Metotları*. Saydam Matbaacılık, Ankara.

Tatar, N., & Buldur, S. (2013). Improving preservice science teachers' self-efficacy about the use of alternative assessment: Implication for theory and practice. *Journal of Baltic Science Education, 12*(4), 452.

Üztemur, S. S., & Metin, C. (2015). Sosyal bilgiler öğretmenlerinin ölçme ve değerlendirme alanındaki kavram yanılgıları ve öz yeterlik inançlarının incelenmesi. *AJELI-Anatolian Journal of Educational Leadership and Instruction, 3*(2), 41-67.

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

Yayla, G. (2011). Fen ve teknoloji öğretmenlerinin tecrübeleriyle alternatif ölçme değerlendirme yaklaşımlarına yönelik öz yeterlilikleri arasındaki ilişki. *International Conference on New Trends in Education and Their Implications,* Siyasal Kitabevi, 879-883.