

An Analysis of Self-, Peer-, and Teacher- Assessment within the Scope of Classroom Teaching Activities

OPEN ACCESS

Volume: 9

Special Issue: 2

Month: September

Year: 2021

E-ISSN: 2582-1334

Received: 10.07.2021

Accepted: 20.08.2021

Published: 15.09.2021

Citation:

Cirit, Didem Karakaya. "An Analysis of Self-, Peer-, and Teacher- Assessment within the Scope of Classroom Teaching Activities." *Shanlax International Journal or Education*, vol. 9, no. S2, 2021, pp. 150–63.

DOI:

<https://doi.org/10.34293/education.v9iS2-Sep.4380>



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

Didem Karakaya Cirit

Munzur University, Turkey

 <https://orcid.org/0000-0002-8606-478X>

Abstract

The aim of this study was to determine the correlation between self-, peer- and teacher- assessment to evaluate preservice science teachers' classroom teaching activities. A mixed method was employed. The sample consisted of 55 senior students (29 women, 26 men) from the science teaching program of a public university in Turkey. Quantitative data were collected using a classroom observation form, which was the Reformed Teaching Observation Practice (RTOP), while qualitative data were collected using observation notes. The study was conducted within the scope of the course "applied teaching" for three weeks under the scope of three topics; global warming (GW), acid rain (AR), and ozone depletion (OD). Each participant attended nine assessment processes with two peers for the three topics. Quantitative results did not show a correlation between self- and teacher-assessment on the three topics. There was no correlation between GW self-assessment and GW peer-assessment and between AR peer-assessment and OD peer-assessment. However, there was a correlation between OD and AR self- and peer-assessment. There was a correlation between peer-assessment and teacher-assessment on neither of the three topics. Qualitative results showed that participants with high RTOP scores in peer-assessment were more likely to make quite superficial qualitative assessments, and briefly describe the teaching process and positively assess it. In self-assessment, participants not only gave themselves high scores but also positively described the teaching process. In teacher-assessment, quantitative and qualitative assessment was consistent.

Keywords: Assessment, Science education, Classroom teaching

Introduction

Alternative educational approaches and methods that help students develop 21st century skills have become popular in recent years. For example, STEM (science, technology, engineering, and mathematics) education has been proven to be effective for 21st century skills and been pursued internationally since the mid-2000s. Teachers play a key role in helping students develop 21st century skills. Only innovative and equipped teachers who can make effective decisions, solve problems, and recognize their own potential and strengths and weaknesses can help students develop those skills. Ideal teachers are those who know their students well, recognize their potential, use appropriate methods and approaches for determining students' levels, know how to use the necessary approaches and means for assessing what students learn and how they learn it, and integrate new techniques and technologies into their classes. The "Pedagogical Content Knowledge (PCK)" model proposed by Shulman (1986, 1987) to define the knowledge fields that equipped learners should have also addresses those skills. In the following years, new ones have been added, while some others have been removed from the PCK. Technology has become more and more prominent in the field of education, leading to changes in the definitions of skills, hence, the emergence of Technological Pedagogical Content Knowledge (TPCK).

The TPCK is defined by Mishra and Koehler (2006) as a new field of knowledge integrating pedagogy, technology, and content knowledge. Every field of knowledge is essential, but assessment plays a crucial role in promoting students' learning and encouraging them to develop 21st century skills and to take responsibility for their own learning, and judge and reflect it. Therefore, teachers who know assessment approaches well and use them effectively are more likely to involve their students in assessment processes and teach them how to use them.

Assessment is used to find out about the strengths and weaknesses of the learning process and students' development (Pandra & Mardapi, 2017). There is a consensus among researchers that multi-directional assessment is more efficient than conventional one-directional assessment, which is however predominantly used in current educational settings (Orsmond, Merry & Reiling, 200; Pope, 2005). Conventional assessment methods (written exams, filling in the blanks, multiple choice tests, true-false, etc.) have lost their popularity in recent years because they are unable to focus on students' responsibilities (Jafarpur1991; McNamara, 2001). In recent years, assessment methods involving students in the process have become more and more popular. The active role of students in assessment is examined in two parts: 1-self-assessment and 2- peer-assessment. Self-assessment and peer-assessment attract more and more attention to achieve effective learning (Wanner & Palmer, 2018). Researchers also argue that self-assessment and peer-assessment improves the quality of learning (Dochy, Segers & Sluijman, 1999; Poon, McNaught, Lam & Kwan, 2009). For self-assessment, every student should think honestly and critically about their own performance (Bozkurt, 2020). Self-assessment helps students identify the knowledge or skills that they need and judge their own learning (performance and achievement) (Boud & Falchikov, 1989). Self-assessment is a basic skill for self-regulatory and lifelong learning (Boud, 1995; Kirby and Downs, 2007; Tan, 2012). Peer assessment, on the other hand, involves peer feedback that can make students more motivated (Chen, 2010). Some studies show that peer-assessment promotes students' learning (Ballantyne, Hughes & Mylonas, 2002). However, students

consider peer-assessment challenging (Falchikov, 1986; Kearney, 2013). Students' negative attitudes and resistance to self- and peer-assessment are a great challenge for successful practices (Kaufman & Schunn, 2011; Van Zundert, Sulijmans & Van Merriënboer, 2010). However, peer-assessment improves students' learning and encourages them to embrace the assessment process (Bryant & Carless, 2010, p.3).

Literature Review

Research on assessment mostly addresses English as a foreign language (EFL) (Birjandi & Hadidi Tamjid, 2012; Cheng, Rogers & Wang, 2008; Cumming, 2001; Khonbi & Sadeghi, 2013; Matsuna, 2009; Nejad & Mahfoodh, 2019; Sadeghi & Khonbi, 2015; Wei, 2015). For example, Sadeghi and Khonbi (2015) investigated university students' attitudes towards self-, peer-, and teacher-assessment and their impact on academic performance. The researchers reported that peer-assessment had a positive effect on students' academic performance. Research generally focuses on improving preservice teachers' assessment skills (Seifert & Feliks, 2019), anxiety about self- and peer-assessment (Gurbanov, 2016; Nawas, 2020; Pope, 2005), preservice teachers' self and assessment (Bozkurt, 2020; Lynch, McNamara, & Seery, 2012), comparison of self- and peer-, and teacher-assessment (Tait-McCutcheon & Knewstubb, 2018), and technology-based assessment (Lin, Liu, & Yuan, 2001; Seifert & Feliks, 2018; Sung, Chang, Chiou, & Hou, 2005; Zheng, Cui, Li, & Huang, 2018). For example, Zheng et al. (2018) investigated whether simultaneous discussion between peer-evaluators and self-evaluators in a web-based assessment system improved peer feedback, students' writing performance, metacognitive awareness, and self-efficacy. Their survey data showed that simultaneous discussion helped improve writing performance, metacognitive awareness, and self-efficacy, and their interview data showed that simultaneous discussion allowed evaluators to give more emotional and metacognitive feedback and reflect it on their writing. Bozkurt (2020) employed a phenomenological method to determine 21 preservice teachers' views of self- and peer-assessment and conducted semi-structured interviews and took observation notes.

He concluded that self- and peer-assessment was not only an assessment tool but could also be used as a powerful learning activity. Tait-McCutcheon and Knewstubb (2018) conducted a case study to compare the self-, peer- and teacher-assessment dynamics among 34 preservice teachers. Peers and teachers assessed the participants' products using an assessment rubric and a feedback chart. The researchers found that more than half (59%) of the participants had consistent self-, peer-, and teacher-assessment results but that the remaining peers and teachers rated the participants' performance lower than the participants rated themselves.

Significance of the Research

Self-, peer-, and teacher-assessment is important for preservice teachers because they should be able to reflect on their own teaching activities (Collin, Karsenti & Komis, 2013). Research looks into students' perceptions of self- and peer-assessment (Mulder, Pearce & Baik, 2014; Van Zundert et al. 2010; Vickerman, 2009). There is a gap in the literature on this issue (Wanner & Palmer, 2018), and therefore, this is the first study to address peer-, self-, and teacher-assessment in a real classroom environment. Most studies assess activities within the scope of courses and outside actual classroom settings (Kılıç, 2016; Nejad & Mahfoodh, 2019; Saito & Fujita, 2009). The second significance of the study is that it focused on assessment under three topics and that preservice teachers who took part in self-assessment also took part in peer-assessment, which is also an understudied issue in the literature. Participants who took part in both self- and peer-assessment had the opportunity to see their strengths and weaknesses both by themselves and with the help of their peers and teachers. Some studies are quantitative (Honsa, 2013; Munoz & Alvarez, 2007; Panadero, Tapia, & Huertas, 2012; Ross, 2006; Sun, Harris, Walther, & Baiocchi, 2015; White, 2009; Willey & Gardner, 2009), while others are qualitative (Azarnoosh, 2013; Harris & Brown 2013; Li & Chen, 2016; Nortcliffe, 2012; Siow, 2015). Therefore, the final significance of this study is that it employed a mixed method by which both qualitative and quantitative data were collected together within the scope of classroom teaching activities. Mixed-

design studies have limitations. The participants rated both themselves and their peers using an observation form and evaluated both themselves and their peers qualitatively. Teachers also took part in this process.

Research Questions

Is there a correlation between self-, peer-, and teacher-assessment of preservice science teachers' classroom teaching about the topics of global warming (GW), acid rain (AR), and ozone depletion (OD)?

1. What kind of correlation is there between self- and peer-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?
2. What kind of correlation is there between self- and teacher-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?
3. What kind of correlation is there between peer- and teacher-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?
4. What are the levels of the pre-service teachers' self, peer, and teacher-assessment of classroom teaching in three science subjects?
5. How is the self-, peer-, and teacher-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?

Method

The aim of this study was to determine the correlation between the self-, peer-, and teacher-assessment of preservice science teachers' (PST) classroom teaching activities. A mixed method of qualitative and quantitative data collection was employed (Creswell & Plano-Clark, 2011). Mixed design involves the collection, analysis, and interpretation of qualitative and quantitative data in a single study (Onwuegbuzia & Leech, 2006, p.474).

Participants

The sample consisted of 55 senior students (29 women, 26 men) from the science teaching program of a public university in Turkey. Participants were recruited using convenience sampling.

Research Process

The study was conducted within the scope of the course “applied teaching.” Participants lectured the course “applied teaching” in an actual middle-school classroom for three weeks. In the first week, they took part in one self- and two peer-assessment sessions within the scope of the topic of GW. In other words, each participant took part in three assessment sessions in a week and was assessed by their teacher. The same procedure was carried out for the topics of AR and OD. Each participant assessed their two peers about each topic and assessed the same peers about each topic. For peer-assessment, the participant completed the Reformed Teaching Observation Practice (RTOP) form and took observation notes while watching her peer perform teaching activities. She then qualitatively described her peer’s performance in detail. For self-assessment, the participant performed her teaching activities and then completed the RTOP form by herself without interacting with anyone, and then, noted down a few more things on the observation notes section. For teacher-assessment, the teacher completed the Reformed Teaching Observation Practice (RTOP) form and took observation notes

while watching the participant perform teaching activities. Figure 1 shows the relationship between self-, peer-, and teacher-assessment. To ensure objectivity, participants were informed that they would not be graded on their performance and that there was no pass/fail associated with the assessment process.

Figure 1: Relationship between Self-, Peer, and Teacher-Assessment



Data Collection

Qualitative and quantitative data collection tools were used together. Quantitative data were collected using the RTOP protocol, while qualitative data were collected using observation notes and video-records of the classroom teaching. A total of 440 qualitative and quantitative data were collected on the topic of GW. The same procedure was applied for the topics of AR and OD (table 1).

Table 1: Data Collection Tool and Distribution of Self-, Peer-, and Teacher-Assessment

Topics	Self-Assessment			Peer-Assessment			Teacher-Assessment		
	RTOP	Observation Note	Total	RTOP	Observation Note	Total	RTOP	Observation Note	Total
Global Warming	55	55	110	110	110	220	55	55	110
Acid Rain	55	55	110	110	110	220	55	55	110
Ozone Depletion	55	55	110	110	110	220	55	55	110

Reformed Teaching Observation Practice

The Reformed Teaching Observation Practice (RTOP) protocol was used to determine the self-, peer-, and teacher- assessment of participants’ classroom teaching on global issues. Inside the classroom: Observation and Analytical Protocol (Horizon Research, 2000), General Information Configuration Model - Classroom Observation Protocol (Ebenezer, et al. 2010), Holistic learning environment (Keser, 2003), and RTOP (Piburn, et al. 2002) on science education were examined

in accordance with the purpose of this study. The RTOP was the protocol of choice because it is an observation instrument for determining to the degree to which science and mathematics teachers perform classroom teaching. The RTOP consists of 25 items scored on a 5-point Likert scale (0 = behavior was never observed to 4 = very descriptive of observed behavior), with the total score ranging from 0 to 100.

Observation Notes

The peers and teacher observed the participant give lectures on the phenomena of GW, AR, and OD in an actual middle school classroom. The researcher was also an observer. In this approach, the items of a particular protocol are not marked, but on the contrary, what is important is that the observer knows what to pay attention to concerning the research topic (Mayring, 2000). The goal of observation notes is to qualitatively analyze self-, peer-, and teacher-assessment. The observation notes in this study were used to identify how aware the participant and her peers and teacher was of her strengths and weaknesses of her performance.

Data Analysis

The observation notes data were analyzed using content analysis. Pearson Correlation analysis was used in determining the relationships between pre-service teachers' self-, peer-, and teacher- assessment.

Results

This section addressed the results of the self-, peer-, and teacher- assessment of participants' teaching performance on global environmental issues.

1. What kind of correlation is there between self- and teacher-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?

There was a significant correlation between self and teacher-assessment in neither of the three topics (Table 2). There was a significant correlation only between OD self-assessment and AR teacher-assessment ($r = 0.271, p = 0.046$). There was no significant correlation between GW self- and teacher-assessment ($r = 0.110, p = 0.422$). There was no significant correlation between GW self-assessment and AR teacher-assessment ($r = 0.143, p = 0.298$). There was also no significant correlation between GW self-assessment and OD teacher-assessment ($r = -0.120, p = 0.383$).

Table 2: Self- and Teacher-Assessment of Participants' Performance of Teaching Global Warming, Acid Rain, and Ozone Depletion

Components	GWSA	GWTA	ODSA	ODTA	ARSA	ARTA
Global Warming Self-Assessment(GWSA)	-					
Global Warming Teacher-Assessment(GWTA)	.110	-				
Ozone Depletion Self-Assessment(ODSA)	.495**	.053	-			
Ozone Depletion Teacher-Assessment(ODTA)	-.120	.127	-.127	-		
Acid Rain Self-Assessment (ARSA)	.469**	.095	.720**	-.167	-	
Acid Rain Teacher-Assessment(ARTA)	.143	.097	.271*	.118	.111	-

$p < .05$; $**p < .01$

There was no significant correlation between OD self- and teacher-assessment ($r = -0.127, p = 0.357$). There was no significant correlation between OD self-assessment and GW teacher-assessment ($r = 0.053, p = 0.700$). There was no significant correlation between AR self- and teacher-assessment ($r = 0.111, p = 0.419$). There was no significant correlation between AR self-assessment and GW teacher-assessment ($r = 0.095, p = 0.489$). Lastly, there was no significant correlation between AR self-

assessment and OD teacher-assessment ($r = -0.167, p = 0.222$).

2. What kind of correlation is there between self- and peer-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?

Table 3 shows the correlation between GW, AR, and OD self- and peer-assessment. There was no significant correlation between GW self- and peer-assessment ($r = 0.208, p = 0.127$). There was no

significant correlation between GW self-assessment and AR peer-assessment ($r = 0.125$, $p = 0.362$). There was no significant correlation between GW self-assessment and OD peer-assessment ($r = 0.114$, $p = 0.408$).

Table 3. Self- and Peer-Assessment of Participants' Performance of Teaching Global Warming, Acid Rain and Ozone Depletion

Components	GWSA	ODSA	ARSA	ARPA	GWPA	ODPA
GWSA	-					
ODSA	.495**	-				
ARSA	.469**	.720**	-			
ARPA	.125	.459**	.633**	-		
GWPA	.208	.320*	.552**	.581**	-	
ODPA	.114	.498**	.588**	.738**	.521**	-

$p^* < .05$; $**p < .01$

There was a correlation between AR self- and peer-assessment ($r = 0.633$, $p = 0.000$). There was a correlation between AR self-assessment and GW peer-assessment ($r = 0.552$, $p = 0.000$). There was also a correlation between AR self-assessment and OD peer-assessment ($r = 0.588$, $p = 0.000$). There was a correlation between OD self- and peer-assessment ($r = 0.0498$, $p = 0.000$). There was a correlation between OD self-assessment and AR peer-assessment ($r = 0.459$, $p = 0.000$). There was a correlation between OD self-assessment and GW peer-assessment ($r = 0.320$, $p = 0.017$) (Table 3).

3. What kind of correlation is there between peer- and teacher-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?

Table 4 shows the correlation between GW, AR, and OD peer- and teacher-assessment. There was no significant correlation between GW peer- and teacher-assessment ($r = 0.058$, $p = 0.674$). GW peer-assessment was not significantly correlated with AR teacher-assessment ($r = 0.116$, $p = 0.400$) and OD teacher-assessment ($r = -0.101$, $p = 0.941$).

Table 4: Peer- and Teacher-Assessment of Participants' Performance of Teaching Global Warming, Acid Rain, and Ozone Depletion

Components	GWPA	ODPA	ARPA	GWTA	ODTA	ARTA
GWPA	-					
ODPA	.521**	-				
ARPA	.581**	.738**	-			
GWTA	.058	-.139	.083	-		
ODTA	-.010	-.113	.078	.127	-	
ARTA	.116	.223	.161	.356	.118	-

$**p < .01$

There was no correlation between AR peer and teacher-assessment ($r = 0.161$, $p = 0.240$). AR peer-assessment was not significantly correlated with GW teacher-assessment ($r = 0.083$, $p = 0.546$) and OD teacher-assessment ($r = 0.078$, $p = 0.570$). There was no correlation between OD peer- and teacher-assessment ($r = -0.113$, $p = 0.411$). There was also no correlation between OD peer-assessment and GW teacher-assessment ($r = -0.139$, $p = 0.310$). There

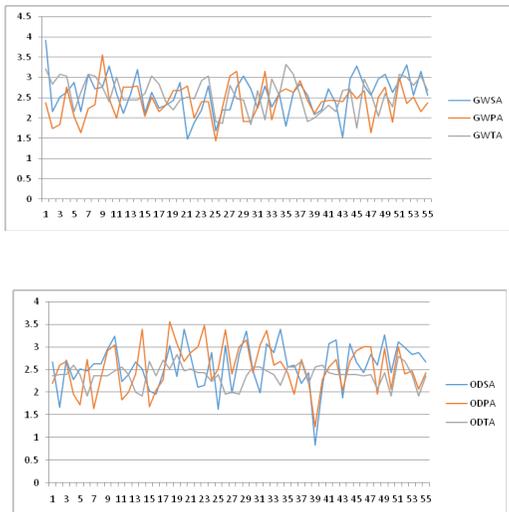
was no significant correlation between OD peer-assessment and AR teacher-assessment ($r = 0.223$, $p = 0.102$).

4. What are the levels of the pre-service teachers' self, peer, and teacher-assessment of classroom teaching in three science subjects?

There was a difference between GW self- and teacher-assessment (Figure 2). The mean GW self- and teacher-assessment scores were 1.8 and 3.32,

respectively, (Figure 2), suggesting that participants scored their GW teaching performance lower than the teacher did. The opposite was also the case. For example, participant 1 had a mean self- and teacher-assessment score of 3.92 and 3.2, respectively. However, the GW teacher-assessment scores were overall higher than the self-assessment scores. There was a difference between the mean GW self- and peer-assessment scores. For example, participant 21 had a mean GW self- and peer-assessment score of 1.48 and 2.8, respectively. On the contrary, participant 45 had a mean GW self- and peer-assessment score of 3.28 and 2.48, respectively. There was a difference between the GW peer- and teacher-assessment scores. Most participants had GW peer-assessment scores of 2 to 3.5. However, some participants had scores below that range. For example, participant 6 had a mean GW peer-assessment score of 1.64 (Figure 2). The mean GW teacher-assessment scores ranged from 2 to 3.5. However, some participants had GW teacher-assessment scores below that range. For example, participant 45 had a mean GW teacher-assessment score of 1.76.

Figure 2: Self- Peer and Teacher-Assessment of Participants' Performance of Teaching Global Warming, Acid Rain, and Ozone Depletion



AR teacher-assessment scores were higher than self-assessment scores (Figure 2). Participants' self-assessment scores were concentrated in the range of 2 to 3. However, some participants had higher

or lower self-assessment scores than that range. For example, participant 14 had a mean self-assessment score of 3.84, whereas participant 50 had a mean self-assessment score of 1.68. Participants had similar mean AR self- and peer-assessment scores (Figure 2), and some even had the same self and peer-assessment scores. For example, participant 9 had a mean self and peer-assessment score of 3.2. Figure 2 shows the mean AR peer- and teacher-assessment scores. Participants had a mean AR peer-assessment score of 3.5 to 1.5. However, participant 39 had a mean AR peer-assessment score of 1.48. Participants had a mean AR teacher-assessment score of 3.5 to 2.

Most participants had a mean OD self-assessment score of 2 to 3, while some others had higher or lower scores than that range. For example, participant 39 had a mean self-assessment score of 0.84, whereas participant 34 had a mean self-assessment score of 3.4. Overall, participants had a higher mean OD self-assessment score than teacher-assessment score, which ranged from 2 to 2.5. Participants also had similar mean OD self- and peer-assessment scores. For example, participant 13 had a mean self and peer-assessment score of 2.68 and 2.4, respectively. They had a mean OD peer- and teacher-assessment score of 3.5 to 1.5 and 2 to 2.5, respectively. However, some participants had lower OD teacher-assessment scores than that range. For example, participant 51 had a mean OD teacher-assessment score of 2.8.

5. How is the self-, peer-, and teacher-assessment of preservice science teachers' classroom teaching about the topics of GW, AR, and OD?

Most participants described what they did during class and talked about their strengths. They mostly focused on the introduction to the lesson, the type of activities, the use of technology, and stages of assessment, and talked about what they did or did not at those stages. Participants explicitly described the classroom teaching they observed, highlighted the strengths of their peers more than their weaknesses, but emphasized little what was missing, wrong, or inadequate during the lectures. Participants evaluated their peers' performance as "It was not a bad class," "It was a good class," "The videos were good," and "The assessment was good." However, some participants evaluated their peers incompetently

or incorrectly. They made such statements as “The teacher played a video which was not very clear,” “The teacher confused the students about the concept of ozone layer,” “The teacher was not able to fully motivate the students,” “The teacher just kept talking for 25 minutes,” “I didn’t like the way the teacher motivated the students,” “The teacher should have played a different video,” “I don’t think that the teacher made the learning outcome clear enough,” “The students’ ideas were not contextualized,” “The introduction was a bit weak,” and “The introduction was dull.”

Participants with high self- and peer-assessment scores considered the teaching process ineffective and found it time-consuming to discuss students’ ideas and stated that they had fallen behind the course material due poor time management. For example, participant 6 noted that she had fallen behind the course material and could not thoroughly address the concept of greenhouse effect. Participant 16 remarked that she could not dispel her students’ misconception that global warming was caused by the hole in the ozone layer. Both participants had the highest scores on Items 6 and 7 of the RTOP form that concerned with conceptual knowledge. Item 6 was “The class covered the basic concepts of the topic,” while Item 7 was “The class promoted meaningful learning.” In their qualitative assessment, the teacher highlighted participants’ strengths and weaknesses regarding their classroom teaching and critically evaluated the teaching process rather than simply describing it. Moreover, the quantitative and qualitative assessment scores given by the teacher were consistent. In other words, the teacher qualitatively justified why she gave high or low assessment scores to participants’ performance. Examples of GW self-, peer-, and teacher- assessment are presented below.

“The participant presented the average global temperature rise per year at the beginning of the class, which was good, but at first the students were not interested in it, so she involved the students in discussion on the topic. Some students were interested, but they were mostly speaking too loudly, which disrupted the flow of the class. The participant performed well, she made sure that the students understood the topic. She explained the concepts of global warming. She got the students to come up

with slogans, which was good for assessment.” (Peer Assessment, Participant 30)

“I tried to use the 5E model. I told them the average temperature in Turkey ten years ago and this year to determine priorknowledge. I asked the students about the possible causes and consequences of the difference. However, the students were not interested at all, and kept talking to each other. I had a very hard time getting their attention. In the discovery part, I played a video about the consequences of global warming, and then, we had discussion about them. Some students said that global warming was not that bad because they could put on summer clothes all the time. So, I asked them the downsides of global warming, and we discussed them, and we agreed that it had too many downsides. In the expansion part, I asked them what to do to stop global warming. They said such things as we should protect the nature and should not waste water. They didn’t know that electric vehicles had an impact on global warming. I divided the class into groups for assessment. I told them to imagine that they were environmentalist groups and asked them to write slogans to draw attention to global warming and write down what to do to stop it. They came up with seven slogans. I told them we could vote to choose the best one, but they were making too much noise, and so I collected the papers and finished the class...As for the global warming topic, I guess I just couldn’t explain the greenhouse effect and ozone depletion, I mean, it would have been better if I had dwelled on them a little bit more, but I couldn’t do it because I was supposed to teach them the consequences of global warming and what measures to take” (Self-Assessment, Participant 33).

“The participant started the class by drawing on the blackboard the temperature rises in Turkey between 1999 and 2010, but instead she should have made an introduction with an incident that would attract students’ attention. Then she posed knowledge-focused questions rather than questions that would get the students to question things. She tried to set a discussion-inducing environment, but it was mostly in the form of Q&A. She did not pose such questions as “Do electrical devices cause global warming?” or “How do you think radiation may affect global warming? to get the students

to question things. She also made statements that might cause misconceptions. For example, she said “radiation causes a hole in the ozone layer, which results in global warming.” It was a mistake on her part that she tried to set a discussion-inducing environment without first determining the students’ prior knowledge. Instead, she should have used concept maps, drawings, etc. to determine their prior knowledge. She was unable to control the class and had a hard time managing it. She was merely explaining things throughout the lesson. She was unable to provide an inquiry-inducing environment for students. The assessment process was an effective and innovative approach for active student engagement, but she had a hard time managing it too, because the students were talking too much.” (Teacher-Assessment)

Conclusion and Discussion

The results can be summarized under three headings: The first is the correlation between self-, peer-, and teacher- assessment. Self- and peer-assessment is of particular importance for preservice teachers because it helps teachers develop self-regulation, critical thinking, and problem-solving skills (Kılıç, 2016). The results indicated a correlation between self- and teacher-assessment in none of the topics (Table 2). Some other studies show that self- and teacher-assessment is consistent (Bouzidi & Jailliet, 2009; Cho, Schunn & Wilson, 2006; Sadler & Good, 2006; Tsai & Liang, 2009; Tseng & Tsai, 2007). However, the results suggested a correlation between OD self-assessment and AR teacher-assessment. Earlier studies have also found a high correlation between self- and teacher-assessment (Cho, Schunn, & Wilson, 2006; Tsai & Liang, 2009; Tseng & Tsai, 2007). GW self-assessment was not correlated with GW peer-assessment, and AR peer-assessment was not correlated with OD peer-assessment. However, OD self-assessment was correlated with OD peer-assessment, AR self-assessment was correlated with AR peer-assessment (Table 3). OD and AR self-assessment was also correlated with GW peer-assessment. The difference in the results concerning global warming may be because it was the first time when participants lectured on it and were involved in the assessment process, which was GW,

AR, and OD in the first, second, and third weeks, respectively. Therefore, they probably familiarized themselves with the assessment process in the first week. It may also be because they had been informed that they would not be graded on their performance and that there was no pass/fail associated with the assessment process. If peer-assessment is not based on careful training, then the results may be based on friendship rather than learning outcomes (Dochy, Segersi & Sluijsmans, 1999). There was a correlation between peer-assessment and teacher- assessment in neither of the three topics. Chen (2010) also reported inconsistency between peer- and teacher-assessment.

The second result is the mean self-, peer-, and teacher-assessment scores. There was a difference between the mean GW self- and teacher- assessment scores. Participants’ mean GW self-assessment score was lower than their teacher-assessment score. Participants had a higher mean AR teacher-assessment score than AR self-assessment score (Figure 2). Tait-McCutcheon and Knewstubb (2018) also reported lower self-assessment scores than peer- and teacher-assessment scores. However, participants had higher OD self-assessment scores than teacher-assessment scores. This is probably because they got used to the process and had higher confidence and better judgment skills. Research already shows that most students find the assessment process challenging, time-consuming, and socially inappropriate. However, assessment plays a critical role in improving the quality of learning outcomes and developing skills and abilities (Topping, Smith, Swanson, & Elliot, 2000). There was a difference between the mean GW self- and peer-assessment scores. Limited education on the purpose and principles of self- and peer-assessment can cause problems (Sullivan and Hall, 1997; Kearney, 2013). The mean AR and OD self- and peer-assessment scores were similar, but the former was slightly lower than the latter. Rudy, Fejfar, Griffith, and Wilson (2001), on the other hand, reported higher peer-assessment scores than self-assessment scores. The mean peer and teacher-assessment scores differed on all three topics. The mean GW and AR teacher- and peer-assessment scores differed, but the former was slightly higher than the latter, which was the exact opposite of the mean OD teacher- and peer-

assessment scores. Participants had a higher mean OD peer-assessment score than teacher-assessment score, which has been reported by earlier studies as well. Kılıç (2016) had preservice teachers make 30-minute presentations on any topic within the scope of the course “Teaching Principles and Methods.” The presentations were then assessed by participants, and their peers and teacher. The results showed that the peer-assessment scores were higher than the self- and teacher-assessment scores. Some studies, on the other hand, report higher peer-assessment scores than teacher-assessment scores (Magin & Helmore, 2001; Rudy, et al. 2001), whereas some others report similar peer- and teacher-assessment scores (Falcikov, 1995; Freeman, 1995; Stefani, 1994).

The third result is the qualitative assessment of the participants. The qualitative results showed that participants with high RTOP scores in peer-assessment were more likely to make quite superficial qualitative assessments and briefly describe the process and positively assess it. Some participants (N = 35%) gave their peers high RTOP scores, but in their qualitative assessment, they addressed mostly their weaknesses and explained in a long and detailed way that their peers were unable to perform the teaching process effectively. In self-assessment, participants gave themselves high RTOP scores and described their teaching performance positively. Some participants (15%) gave themselves low RTOP scores and evaluated their performance poorly. The quantitative and qualitative teacher-assessment was consistent. The science teaching undergraduate curriculum in Turkey consists of a total of 64 courses, 16 of which are elective. Of those 64 courses in the 2018-renewed undergraduate curriculum of the Council of Higher Education (CHE), 22 are vocational knowledge, 12 general culture, and 30 content knowledge (CHE, 2018). First-year students mostly take field (physics, chemistry, and biology) and general cultural knowledge courses (foreign language, information technologies, etc.), while second- and third-year students mostly take content knowledge courses. Fourth-year students take vocational courses (measurement and assessment in education, research methods in education, etc.) and content knowledge courses, but not general culture courses. Of these courses, the “Measurement

and Assessment” is offered in the first semester of the third year. Moreover, faculty exams are based on conventional assessment methods (open-ended questions, multiple choice, or short answer etc.). This may account for why preservice teachers have inadequate self- and peer-assessment skills. Therefore, they should be encouraged to be involved in assessment processes as soon as possible during their undergraduate education, and courses should also adhere to that approach. Feedback should be used in this process, and longer action or experimental research (e.g. 10 weeks) should be conducted.

References

- Azarnoosh, Maryam. “Peer Assessment in an EFL Context: Attitudes and Friendship Bias.” *Language Testing in Asia*, vol. 3, no. 1, 2013, pp. 1-10.
- Ballantyne, Roy, et al. “Developing Procedures for Implementing Peer Assessment in Large Classes using an Action Research Process.” *Assessment & Evaluation in Higher Education*, vol. 27, no. 5, 2002, pp. 427-441.
- Birjandi, Parviz, and Nasrin Hadidi Tamjid. “The Role of Self-, Peer and Teacher Assessment in Promoting Iranian EFL Learners’ Writing Performance.” *Assessment & Evaluation in Higher Education*, vol. 37, no. 5, 2012, pp. 513-533.
- Boud, David, and Nancy Falchikov. “Quantitative Studies of Student Self-Assessment in Higher Education: A Critical Analysis of Findings.” *Higher Education*, vol. 18, no. 5, 1989, pp. 529-549.
- Boud, David. *Enhancing Learning through Self-Assessment*. Routledge, 1995.
- Bozkurt, Fisun. “Teacher Candidates’ Views on Self and Peer Assessment as a Tool for Student Development.” *Australian Journal of Teacher Education*, vol. 45, no. 1, 2020, pp. 47-60.
- Bouzidi, L’hadi, and Alain Jaillet. “Can Online Peer Assessment Be Trusted?.” *Journal of Educational Technology & Society*, vol. 12, no. 4, 2009, pp. 257-268.
- Bryant, Darren A., and David R. Carless. “Peer Assessment in a Test-Dominated Setting: Empowering, Boring or Facilitating

- Examination Preparation?." *Educational Research for Policy and Practice*, vol. 9, no. 1, 2010, pp. 3-15.
- Chen, Chao-hsiu. "The Implementation and Evaluation of a Mobile Self-and Peer-Assessment System." *Computers & Education*, vol. 55, no. 1, 2010, pp. 229-236.
- Cho, Kwangsu, et al. "Validity and Reliability of Scaffolded Peer Assessment of Writing from Instructor and Student Perspectives." *Journal of Educational Psychology*, vol. 98, no. 4, 2006, pp. 891-901.
- Collin, Simon, et al. "Reflective Practice in Initial Teacher Training: Critiques and Perspectives." *Reflective Practice*, vol. 14, no. 1, 2013, pp. 104-117.
- Creswell, John W., and Vicki L. Plano Clark. "Choosing a Mixed Methods Design." *Designing and Conducting Mixed Methods Research*, vol. 2, 2011, pp. 53-106.
- Cumming, Alister. "ESI/EFI Instructors' Practices for Writing Assessment: Specific Purposes or General Purposes?." *Language Testing*, vol. 18, no. 2, 2001, pp. 207-224.
- Dochy, F., et al. "The Use of Self-, Peer and Co-Assessment in Higher Education: A Review." *Studies in Higher Education*, vol. 24, no. 3, 1999, pp. 331-350.
- Falchikov, Nancy. "Peer Feedback Marking: Developing Peer Assessment." *Innovations in Education and Training International*, vol. 32, no. 2, 1995, pp. 175-187.
- Falchikov, Nancy. "Product Comparisons and Process Benefits of Collaborative Peer Group and Self Assessments." *Assessment and Evaluation in Higher Education*, vol. 11, no. 2, 1986, pp. 146-166.
- Freeman, Mark. "Peer Assessment by Groups of Group Work." *Assessment & Evaluation in Higher Education*, vol. 20, no. 3, 1995, pp. 289-300.
- Gurbanov, Elmira. "The Challenge of Grading in Self and Peer-Assessment (Undergraduate Students' and University Teachers' Perspectives)." *Journal of Education in Black Sea Region*, vol. 1, no. 2, 2016, pp. 82-91.
- Harris, Lois R., and Gavin TL Brown. "Opportunities and Obstacles to Consider When Using Peer-And Self-Assessment to Improve Student Learning: Case Studies into Teachers' Implementation." *Teaching and Teacher Education*, vol. 36, 2013, pp. 101-111.
- Honsa, S. "Self-Assessment in EFL Writing: A Study of Intermediate EFL Students at a Thai University." *Voices in Asia Journal*, vol. 1, no. 1, 2013, pp. 34-57.
- "Inside the Classroom Observation and Analytic Protocol." *Horizon*, 2000.
- Jafarpur, Abdoljavad. "Can Naive EFL Learners Estimate Their Own Proficiency?." *Evaluation & Research in Education*, vol. 5, no. 3, 1991, pp. 145-157.
- Kaufman, Julia H., and Christian D. Schunn. "Students' Perceptions about Peer Assessment for Writing: Their Origin and Impact on Revision Work." *Instructional Science*, vol. 39, 3, 2011, pp. 387-406.
- Kearney, Sean. "Improving Engagement: The Use of 'Authentic Self-And Peer-Assessment for Learning' to Enhance the Student Learning Experience." *Assessment & Evaluation in Higher Education*, vol. 38, no. 7, 2013, pp. 875-891.
- Keser, Ömer Faruk. *Fizik Eğitime Yönelik Bütünleştirici Bir Öğrenme Ortamı Tasarımı ve Uygulaması*. Karadeniz Teknik Üniversitesi, 2003.
- Khonbi, Zainab Abolfazli, and Karim Sadeghi. "The Effect of Assessment Type (Self Vs. Peer Vs. Teacher) on Iranian University EFL Students' Course Achievement." *Language Testing in Asia*, vol. 2, no. 4, 2012, pp. 1-28.
- Kilic, Didem. "An Examination of Using Self-, Peer-, and Teacher-Assessment in Higher Education: A Case Study in Teacher Education." *Higher Education Studies*, vol. 6, no. 1, 2016, pp. 136-144.
- Kirby, Nicola F., and Colleen T. Downs. "Self-Assessment and the Disadvantaged Student: Potential for Encouraging self-Regulated Learning?." *Assessment & Evaluation in Higher Education*, vol. 32, no. 4, 2007, pp. 475-494.
- Li, Ying, and Liping Chen. "Peer-and Self-

- Assessment: A Case Study to Improve the Students' Learning Ability." *Journal of Language Teaching and Research*, vol. 7, no. 4, 2016, pp. 780-787.
- Lin, Sunny San-Ju, et al. "Web Based Peer Assessment: Attitude and Achievement." *IEEE Transactions on Education*, vol. 44, no. 2 2001.
- Lin, Sunny SJ, et al. "Web-Based Peer Assessment: Feedback for Students with Various Thinking-Styles." *Journal of Computer Assisted Learning*, vol. 17, no. 4, 2001, pp. 420-432.
- Lynch, Raymond, et al. "Promoting Deep Learning in a Teacher Education Programme through Self-And Peer-Assessment and Feedback." *European Journal of Teacher Education*, vol. 35, no. 2, 2012, pp. 179-197.
- Magin, Douglas, and Phil Helmore. "Peer and Teacher Assessments of Oral Presentation Skills: How Reliable are They?." *Studies in Higher Education*, vol. 26, no. 3, 2001, pp. 287-298.
- Matsuno, Sumie. "Self-, Peer-, and Teacher-Assessments in Japanese University EFL Writing Classrooms." *Language Testing*, vol. 26, no. 1, 2009, pp. 75-100.
- Mayring, Philipp. "Qualitative Content Analysis." *A Companion to Qualitative Research*, vol. 1, no. 2, 2000.
- McNamara, Tim. "Language Assessment as Social Practice: Challenges for Research." *Language Testing*, vol. 18, no. 4, 2001, pp. 333-349.
- Mishra, Punya, and Matthew J. Koehler. "Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge." *Teachers College Record*, vol. 108, no. 6, 2006.
- Mulder, Raoul A., et al. "Peer Review in Higher Education: Student Perceptions Before and After Participation." *Active Learning in Higher Education*, vol. 15, no. 2, 2014, pp. 157-171.
- Munoz, Ana, and Marta E. Álvarez. "Students' Objectivity and Perception of Self Assessment in an EFL Classroom." *The Journal of Asia TEFL*, vol. 4, no. 2, 2007.
- Nawas, Abu. "Grading Anxiety with Self and Peer-Assessment: A Mixed-Method Study in an Indonesian EFL Context." *Issues in Educational Research*, vol. 30, no. 1, 2020.
- Nejad, Ali Mansoori, and Omer Hassan. "Assessment of Oral Presentations: Effectiveness of Self-, Peer-, and Teacher Assessments." *International Journal of Instruction*, vol. 12, no. 3, 2019, pp. 615-632.
- Nortcliff, Anne. "Can Students Assess Themselves and Their Peers?: A Five Year Study." *Student Engagement and Experience Journal*, vol. 1, no. 2, 2012, pp. 1-17.
- Onwuegbuzie, Anthony J., and Nancy L. Leech. "Linking Research Questions to Mixed Methods Data Analysis Procedures." *The Qualitative Report*, vol. 11, no. 3, 2006, pp. 474-498.
- Orsmond, Paul, et al. "The Use of Student Derived Marking Criteria in Peer and Self-Assessment." *Assessment & Evaluation in Higher Education*, vol. 25, no. 1, 2000, pp. 23-38.
- Panadero, Ernesto, et al. "Rubrics and Self-Assessment Scripts Effects on Self-Regulation, Learning and Self-Efficacy in Secondary Education." *Learning and Individual Differences*, vol. 22, no. 6, 2012, pp. 806-813.
- Pandra, Viktor, and Djemari Mardapi. "Development of Mathematics Achievement Test for Third Grade Students at Elementary School in Indonesia." *International Electronic Journal of Mathematics Education*, vol. 12, no. 3, 2017, pp. 769-776.
- Piburn, Michael, et al. "Reformed Teaching Observation Protocol (RTOP) Reference Manual." *Tempe, Arizona: Arizona Collaborative for Excellence in the Preparation of Teachers*, 2000.
- Poon, Wai-Yin, et al. "Improving Assessment Methods In University Science Education With Negotiated Self-And Peer-Assessment." *Assessment in Education: Principles, Policy & Practice*, vol. 16, no. 3, 2009, pp. 331-346.
- Pope, Nigel K. Ll. "The Impact of Stress in Self-And Peer Assessment." *Assessment & Evaluation in Higher Education*, vol. 30, no. 1, 2005, pp.

- 51-63.
- Ross, John A. "The Reliability, Validity, and Utility of Self-Assessment." *Practical Assessment, Research, and Evaluation*, vol. 11, no. 10, 2006, pp. 1-14.
- Rudy, David W., et al. "Self-and Peer Assessment in a First-Year Communication and Interviewing Course." *Evaluation & The Health Professions*, vol. 24, no. 4, 2001, pp. 436-445.
- Sadeghi, Karim, and Zainab Abolfazli Khonbi. "Iranian University Students' Experiences of and Attitudes Towards Alternatives in Assessment." *Assessment & Evaluation in Higher Education*, vol. 40, no. 5, 2015, pp. 641-665.
- Sadler, Philip M., and Eddie Good. "The Impact of Self-and Peer-Grading on Student Learning." *Educational Assessment*, vol. 11, no. 1, 2006, pp. 1-31.
- Saito, Hidetoshi, and Tomoko Fujita. "Peer-Assessing Peers' Contribution to EFL Group Presentations." *RELC Journal*, vol. 40, no. 2, 2009, pp. 149-171.
- Seifert, Tami, and Orna Feliks. "Online Self-Assessment and Peer-Assessment as a Tool to Enhance Student-Teachers' Assessment Skills." *Assessment & Evaluation in Higher Education*, vol. 44, no. 2, 2019, pp. 169-185.
- Shulman, Lee S. "Those Who Understand: Knowledge Growth in Teaching." *Educational Researcher*, vol. 15, no. 2, 1986, pp. 4-14.
- Shulman, Lee. "Knowledge and Teaching: Foundations of the New Reform." *Harvard Educational Review*, vol. 57, no. 1, 1987, pp. 1-23.
- Siow, Lee-Fong. "Students' Perceptions on Self-and Peer Assessment in Enhancing Learning Experience." *Malaysian Online Journal of Educational Sciences*, vol. 3, no. 2, 2015, pp. 21-35.
- Stefani, Lorraine A.J. "Peer, Self And Tutor Assessment: Relative Reliabilities." *Studies in Higher Education*, vol. 19, no. 1, 1994, pp. 69-75.
- Sullivan, Keith, and Cedric Hall. "Introducing Students to Self-Assessment." *Assessment & Evaluation in Higher Education*, vol. 22, no. 3, 1997, pp. 289-305.
- Sun, Dennis L., et al. "Peer Assessment Enhances Student Learning: The Results of a Matched Randomized Crossover Experiment in a College Statistics Class." *PloS One*, vol. 10, no. 12, 2015, pp. e0143177.
- Sung, Yao-Ting, et al. "The Design and Application of a Web-Based Self-and Peer-Assessment System." *Computers & Education*, vol. 45, no. 2, 2005, pp. 187-202.
- Tait-McCutcheon, Sandi, and Bernadette Knewstubb. "Evaluating the Alignment of Self, Peer and Lecture Assessment in an Aotearoa New Zealand Pre-Service Teacher Education Course." *Assessment & Evaluation in Higher Education*, vol. 43, no. 5, 2018, pp. 772-785.
- Tan, Kelvin H. K. "How Teachers Understand and Use Power in Alternative Assessment." *Education Research International*, 2012.
- Topping, Keith J., et al. "Formative Peer Assessment of Academic Writing Between Postgraduate Students." *Assessment & Evaluation in Higher Education*, vol. 25, no. 2, 2000, pp. 149-169.
- Tsai, Chin-Chung, and Jyh-Chong Liang. "The Development of Science Activities Via On-Line Peer Assessment: The Role of Scientific Epistemological Views." *Instructional Science*, vol. 37, no. 3, 2009, pp. 293-310.
- Tseng, Sheng-Chau, and Chin-Chung Tsai. "On-Line Peer Assessment and the Role of the Peer Feedback: A Study of High School Computer Course." *Computers & Education*, vol. 49, no. 4, 2007, pp. 1161-1174.
- Van Zundert, Marjo, et al. "Effective Peer Assessment Processes: Research Findings and Future Directions." *Learning and Instruction*, vol. 20, no. 4, 2010, pp. 270-279.
- Vickerman, Philip. "Student Perspectives on Formative Peer Assessment: An Attempt to Deepen Learning?." *Assessment & Evaluation in Higher Education*, vol. 34, no. 2, 2009, pp. 221-230.
- Wanner, Thomas, and Edward Palmer. "Formative Self-And Peer Assessment for Improved Student Learning: The Crucial Factors of

- Design, Teacher Participation and Feedback.” *Assessment & Evaluation in Higher Education*, vol. 43, no. 7, 2018.
- Wei, Wei. “Using Summative and Formative Assessments to Evaluate EFL Teachers’ Teaching Performance.” *Assessment & Evaluation in Higher Education*, vol. 40, no. 4, 2015, pp. 611-623.
- White, Eddy. “Student Perspectives of Peer Assessment for Learning in a Public Speaking Course.” *Asian EFL Journal*, vol. 33, no. 1, 2009, pp. 1-36.
- Willey, Keith, and Anne Gardner. “Changing Student’s Perceptions of Self and Peer Assessment.” *Research in Engineering Education Symposium - REES*, 2009.
- Zheng, Lanqin, et al. “Synchronous Discussion between Assessors and Assesseees in Web-Based Peer Assessment: Impact on Writing Performance, Feedback Quality, Meta-Cognitive Awareness and Self-Efficacy.” *Assessment & Evaluation in Higher Education*, vol. 43, no. 3, 2018, pp. 500-514.

Author Details

Didem Karakaya Cirit, Munzur University, Turkey, **Email ID:** didemkarakayacirit@munzur.edu.tr