Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi
Cilt 16, Sayı 2, Aralık 2022, sayfa 428-455. ISSN: 1307-6086
Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education
Vol. 16, Issue 2, December 2022, pp. 428-455. ISSN: 1307-6086

Research Article

Prospective Mathematics and Physics Teachers' Experiences of Implementing Peer-Assessment

Olga GKIOKA

Boğaziçi Üniversitesi, North Campus, 34342 Bebek, Istanbul, Türkiye,
olga.gkioka@boun.edu.tr, http://orcid.org/0000-0002-0477-3217

Received : 07.06.2022 Accepted : 15.12.2022

Doi: 10.17522/balikesirnef.1127523

Abstract – The aim of the presented research is to investigate how prospective mathematics and physics teachers have experienced peer-assessment. The study took place within the context of one classroom assessment course designed for pre-service teacher education. Ninety-four (94) prospective teachers in four cohorts participated. They were giving written feedback to peer's coursework, writing reflective journals and participating in semi-structured interviews. The participants reported that grading is more difficult than giving feedback because grades are important as they "count" for students' success. Secondly, during the course, their thinking "shifted" to the idea that giving comments is more crucial than grading, because feedback helped them understand the assessment criteria and subsequently, improve coursework. Based on the findings, the main argument of the paper is about the role of pre-service teacher education for the development of classroom assessment skills and assessment knowledge in teachers.

Key words: assessment criteria, feedback, grading, peer-assessment, pre-service mathematics and physics teacher education.

Introduction

The presented study is heavily influenced by the work and research on classroom assessment by Brookhart and her colleagues (Brookhart, 2011; Brookhart et al., 2016) and on peer-assessment assessment by Black and Wiliam (1998), Brookhart (2017) and Shepard and her colleagues (2005). According to Brookhart and colleagues, through classroom assessment, the teacher should aim to enhance and support student learning. This study has used the term "classroom assessment" to include both formative and summative assessment. Brookhart and McMillan (2020) developed a broad definition of classroom assessment as follows:

Classroom assessment is a process that teachers and students use in collecting, evaluating, and using evidence of student learning for a variety of purposes, including diagnosing student strengths and weaknesses, monitoring student progress toward meeting desired levels of proficiency, assigning grades, providing feedback to students and parents, and enhancing student learning and motivation (Brookhart & McMillan, 2020, p. 4).

Assessment requires not only teachers' but also students' active involvement in assessment when they look at each other's work to identify strengths and weaknesses, to give feedback for improvement (peer-assessment) and then improve their own work (self-assessment). The primary goal is to promote learning and advance learners' responsibility for their own learning and development of metacognitive skills. Within the same framework, Brookhart (2017) has also developed the argument that grading and summative assessment should not mean marking only; grading should also guide and support learning. Hence, teachers need to develop grading practices which focus on learning and aim towards support of learning.

The literature review starts with one definition of peer-assessment and discusses how peer-assessment is related to assessment criteria, feedback and grading. Then, research on peer-assessment has been presented. The literature review concludes with research on teacher assessment practices. It is worth noting that the reviewed research studies were conducted in Turkey, as well as Europe and the US.

Peer-Assessment: Assessment Criteria, Feedback and Grading

During the process of peer-assessment, the intention is that both assessee and assessor to benefit from the process. In peer-assessment, students provide and receive feedback more promptly. Topping (2009) defined peer-assessment as follows:

Peer assessment is an arrangement for learners to consider and specify the level, value, or quality of a product or performance of other equal status learners. Products to be assessed can include writing, oral presentations, portfolios, test performance, or other skilled behaviors (p. 20).

Sadler (1989), in his seminal paper, discussed the three elements required for peer-assessment to promote learning:

• Firstly, students need to develop a clear view of the learning goals and a shared understanding of quality work similar to that of the teacher. In simple learner's words: "Where am I trying to go?"

• Secondly, information is needed about the present state of the learner and the expected one. The learner should be able to compare the current level of performance by using quality criteria. In student's simple terms: "Where am I?"

• Lastly, the learner should be able to take action to close the gap. "How best to get there?"

Sadler (1989) stressed the potential value of peer work both because the communication between students is in a language that students themselves naturally use, and also because students learn from taking the role of teachers and examiners. On such basis, Sadler (1989) strongly argued that self- and peer-assessment are of central importance to formative assessment. At the same time, Sadler argued that students must understand the criteria according to which their work will be judged. In a later paper, Sadler (1998) emphasized the need for the feedback to be understood by the learner and that students should be actively involved in further improvement and learning.

Along the same line, Boud (2000) argued that

Unless students are able to use the feedback to produce improved work, through for example, redoing the same assignment, neither they, nor those giving the feedback will know it has been effective (p. 158).

Feedback is only successful if students use it to improve their performance. As such, feedback needs to identify the quality of the work required and then, it needs to indicate the gap between the current achievement (i.e., strengths) and the desired performance (in relation to assessment criteria and the expected quality). Having discussed the important role of peer-assessment in the service of learning, the focus of the discussion will move to research on peer-assessment.

Research on Peer-Assessment

Research studies on peer-assessment have shown the value and the importance of student involvement in the process. They have provided strong evidence that peer-assessment and good quality feedback can promote learning and encourage progress in student learning in secondary education (i.e., Black & Harrison, 2004; Hodgen & Wiliam, 2006; Ketonen et al., 2020) and in higher education (Sluijsmans & Prins, 2006; Sluijsmans et al., 2002). Wiliam and Leahy (2015) labelled peer-assessment with the expression "activating students as learning sources for one another" and self-assessment with "activating students as owners of their own learning". Peer-assessment helps learners to carry on with the "next" step in the route of their learning, which is self-assessment. Yet, peer-assessment may help students develop ownership of learning when they are actively involved in assessing peer's work (Shepard, 2000).

According to research studies on formative assessment, feedback is an essential component of assessment because it has major influence on learning (Black, 2003; Black & Wiliam, 1998; Hattie, 2011). Shute (2008) used the term "formative feedback" and defined it as

information communicated to the learner that is intended to modify his or her thinking or behaviour for the purpose of improving learning (p. 154).

Butler (1988) showed that students get distracted by grades that go with written comments for improvement. This is because grades make them compare themselves, whilst written comments based on the task make them think about how to improve the work and receive the message that they can do better. In addition, marks only do not give any information about current achievement and necessary improvement. Comments only may support further learning.

Of course one cannot be sure that student peer feedback is always correct as teacher feedback. However, student peer feedback is readily and more immediately available than that of the teacher's. Research undertaken by Topping (2009) provided evidence that once learners become familiar with the process of peer-assessment (after time and adequate practice), reliability is likely to be quite high. Topping (2009) showed that a peer assessor, with less skill in assessment but more time in which to do it, can produce an assessment of equal validity and reliability to that of a teacher. Along the same line, Falchikov and Goldfinch (2000) carried out a meta-analysis of forty-eight quantitative peer assessment studies that compared peer and teacher marks. They concluded that students are able to make reliable judgements because peer-assessments were found to be similar to teacher assessments when students made judgements based on explicit criteria that they had well understood and had taken ownership of them. In Taiwan, Tseng and Tsai (2007) found out that peer assessment grades were highly correlated with teacher scores. Yet, the study by O'Donnell and Topping (1998) provided evidence that peer assessment can be as effective as teacher assessment and sometimes more effective.

Research on Teacher Assessment Practices

While the general research interest in classroom and formative assessment is high, research on teachers' competencies of assessment practices showed that in-service teachers' assessment knowledge and assessment strategies are limited. For example, Campbell and Collins (2007) and DeLuca (2012), based on documentation of teachers' classroom assessment practices, concluded that they are not skillful in assessment. The same researchers made the point that teacher education needs to give more attention to the development of classroom assessment strategies by teachers.

Brookhart and her colleagues (2016) strongly argued that teachers need help to improve the assessment criteria they use to grade, they need to be able to develop assessment criteria and then to effectively incorporate instructional skills with assessment skills. Brookhart (2011), not only argued in favor of the need for educators to give much attention to the development of assessment knowledge and skills for teachers, but she also proposed a list of assessment skills informed by current advances into research on classroom assessment.

Research on peer-assessment in Turkey

In Turkey, Acar-Erdol and Yıldızlı (2018) aimed to identify the classroom assessment practices implemented by instructors at the primary, secondary and high school level. By using Classroom Assessment Practices Survey, they found out that teachers had adopted an approach of assessment for learning and provided strong evidence of the advantages for students' learning. Secondly, Özdemir (2016) worked with prospective teachers to firstly give them the opportunity to practice peer-assessment and then, investigate the participants' views and opinions about peer-assessment. The participants talked extensively about the importance of providing constructive feedback and not grades. Again, from the teachers' perspective, Gelbal and Kelecioğlu (2007) provided evidence of the teachers' need to have practice and initial education in assessment since they identified problems in practice. In addition, Koç (2011) conducted a research study to document prospective teachers' opinions about peer-assessment in teaching practice. The participants talked about a wide range of benefits for both students and themselves as prospective teachers. Ozmen and Aydın (2015) conducted a research study within a teacher education program in Turkey with the aim to examine student teachers' beliefs about oral corrective feedback. Although the subject field was that of language learning and teaching (different from mathematics and physics), the findings are of significant importance for this study. Their evidence showed that although most participants held constructivist beliefs about teaching and learning, their strategies related to corrective oral feedback varied in terms of correcting errors. They reported on such terms as language proficiency, language components and task type.

Sasmaz Oren (2012) investigated the impact of gender and previous experiences on the approach of self- and peer-assessment in Turkey. Interestingly, they found out that female students received significantly higher mean scores than male students. Kayacan and Razı (2017) investigated the impact of self and anonymous peer feedback on four written assignments among high school students. Their findings strongly supported the idea that both self and peer feedback enhance the improvement of assignments in various ways.

In higher education, a study designed and conducted by Şahin (2008) looked at peer-assessment. The aim was to investigate peers when giving scores on project work. It was found that scores by peers were similar to those given by the instructor. In a following study, Şahin and colleagues (2016) looked at peer assessment of undergraduate students to report that those with high levels of achievement differed significantly from those made by students with medium or low level of achievement. More recently, Boztunç Öztürk and colleagues (2019) worked with 66 university students and focused on the analysis and comparison between analytic and general impression scoring in peer-assessment. One main finding was that students were distinguished from one another at a highly reliable rate using various scoring methods.

Method

Based on such research evidence, we, as teacher educators and university teachers, should think about how we develop and provide opportunities for prospective teachers to develop assessment practices in initial teacher education. Consequently, the focus of this research is on how prospective mathematics and physics teachers experienced peer-assessment during initial teacher education. More, specifically, the aim of the study is to explore the difficulties and challenges that prospective mathematics and physics teachers experienced while giving feedback and grades to their peers' coursework.

Research Design and Questions

A qualitative research study was designed to investigate the difficulties and challenges that prospective mathematics and physics teachers experienced while giving feedback and grades to their peers. The approach of the study is that of a case study (Stake, 1995; Yin, 2017) in an effort to collect in-depth and comprehensive information. Stake (1995) defined a case as having specific boundaries in terms of one phenomenon, time and place. For the specific research purposes, the boundaries are the teaching context of the course and the University context within which the research was conducted (as explained in the "Context of the Study" section) .

The study was designed to answer one main research question: "*How did prospective mathematics and physics teachers experience peer-assessment?*" with the following two sub-questions:

1. *What difficulties and challenges did prospective mathematics and physics teachers experience? (when giving feedback and grades to their peers)*

2. How did such experiences change during one academic semester?

It was not within the scope of the study to compare the instructor's feedback and grades with those of the participants. In addition, we are not interested in evaluating the assessment course, within which this research study was conducted. With regard its contribution, the present study aims to highlight the importance of including assessment education and development of assessment knowledge and skills in initial teacher education mainly in initial Mathematics and Physics teacher education.

Participants

Four cohorts of prospective mathematics and physics teachers in their third-year or fourth-year participated at a state university in Turkey. In total, ninety-four (94) undergraduate students were involved (during four consecutive academic semesters: Spring academic semester 2018 - Fall academic semester 2019). They were from the whole range of educational achievement, approximately of equal numbers of mathematics and physics prospective teachers. From the official records, it seems that they were 20-22 years old. In addition, 62% of them were women and 38% men. On average, prospective teachers from the four different academic semesters are samples drawn from the same population. For all of them, Turkish was their native language with English as their first foreign language and official language of teaching at the university. Purposive sampling (Creswell, 2013) was applied, whereby the researcher intentionally selects the participants to investigate the research problem. The names of the prospective teachers have been removed. Instead, we have used numbers for each participant (1-94), as the following Table shows.

Table 1 Cohorts and Participants

	Mathematics	**Physics**	**Total**	**Numbers**
1st Cohort Spring Academic Semester 2018	10	12	22	1-22
2nd Cohort Fall Academic Semester 2018	14	12	26	23-48
3rd Cohort Spring Academic Semester 2019	13	11	24	49-72
4th Cohort Fall Academic Semester 2019	12	10	22	73-94
Total	49	45	94	

The Context of the Study - The Classroom Assessment Course

The study was carried within the context of a classroom assessment course for 13 full weeks, for each of the four cohorts. The course was developed and taught by the main research

investigator, who was the instructor of the four cohorts, too. It was the first compulsory course about classroom assessment for prospective teachers. The participants were introduced to formative and summative assessment, the development of rubrics and grading, feedback and peer-assessment. Further details about the syllabus of the course are given in Appendix A.

The participants were giving oral and written feedback and grades on coursework during one academic semester. Once a week, after the submission of coursework, they would take one peer's piece of coursework as homework (out of class) to give written feedback in the form of comments and grades. Although comments are of major importance, the participants had practice in grading, because grading is a key professional skill for teachers. In addition, one of the main aims of the assessment course was them to develop a sound understanding of grades and grading in support of learning. The instructor would need to look at how well each of them did so that matching of students with coursework was deliberate and selective, according to the achieved coursework. They would bring it back to the class and discuss in pairs for around ten minutes. The participants would talk to each other in groups of two (two different groups for each participant; as assessor and as assessee). For example, if there were 24 prospective teachers in the class, 24 conversations would take place in pairs. Each participant (the assessor, in each pair) gives feedback and they discuss the piece of coursework. Then, they change pairs so that the assessed peer becomes the assessor, who initiates a discussion and gives feedback in a different pair for ten more minutes. They explain the written feedback and they give oral feedback, too. The intention is to identify strengths and strong points, as well as to offer hints for improvement when weak points exist. After peer-assessment, time was given for revision so that each student improved the paper and submitted it again. However, this paper reports only on the peer-assessment process. A different paper has focused on self-assessment experiences by the same prospective teachers. All discussions were recorded with microphones put in front of the participants so that neither the instructor nor the two research assistants interrupted or distorted the natural conversation of students.

Methods of Data Collection

The study employed rigorous data collection involving:

- collection of weekly reflective journals written by the participants,
- semi-structured interviews and,
- collection of written feedback by peers to each other.

Reflective journals (Boud, 2001) were submitted once every week. The participants were asked to write reflective papers in order to talk about their experiences of peer-assessment. They were not asked any particular questions because we would not like them to get biased by our questions. Instead, we wanted them to select the issues they were concerned with (grades, feedback and whether the process helped them learn better). For example, one expression was: "*In your journal, write down how you have experienced peer-assessment this week. What are your experiences when giving and receiving feedback*?" Boud (2001) argued in favour of journals promoting reflection on practice. For Boud, journal writing should capture and enhance reflection on professional practice and learning.

The role of journal writing is to give an account of what happened and to retrieve as fully as possible the rich texture of events as they unfolded (Boud, 2001, p. 14).

Reflective journals were kept private. This gives freedom and confidentiality so that each participant expresses herself and provides a sincere flow of thoughts and feelings. From the beginning of the course, it was clarified that reflective journals were not going to be graded.

Semi-structured individual interviews (Kvale & Brinkmann, 2009) were carried out to shed light on the participants' experiences and difficulties. An interview protocol was developed for the purposes of this study and it was piloted. While the open questions in the list formed the main impetus within the interviews, we were also open to probing ideas that they were raised within the discussion, particularly if the questions seemed important to the participants. Appendix B gives the interview questions. Interviews were carried out individually. Each one lasted approximately for 10-15 minutes. All interviews were audio-taped with the participants' written consent and transcribed. Lastly, written feedback given by the participants was collected for analysis.

Through the three different research methods, experiences are rarely missed but rather collected at different times so that the whole picture can be viewed rather than considering only snapshot examples. Thus, what they have written in reflective journals can be elicited later in the interviews. In addition, the analysis of the feedback (oral and written) they exchange can provide the basis for interviews so that different perspectives can add to the richness of the data. Finally, all participants were informed about the research purposes of the study, gave their written consent and ethics guidelines were kept.

Data Analysis

The analysis of data was an inductive process of narrowing data into a range of emerging themes and trends (Creswell, 2013) through multiple read of the data. Open coding techniques

and triangulation were applied. To make sense out of text data, the text was divided into segments, different segments were identified and the segments were labelled with themes that accurately described the meaning of the text segment. Similar themes were aggregated together to form a major idea until no new themes would emerge (Silverman, 2001). We stopped developing themes when we identified the major themes and no new information could be added to our list of themes or to the detail for existing themes. Thus, themes describing the participants' experiences emerged from the data.

In addition, greater confidence in our findings was developed because the analysis also focused on an extensive use of "triangulation" (Creswell & Miller, 2000) to provide several viewpoints from different participants and sources of data as sufficient evidence for each emerging theme. Triangulation was essential to consistently validate and verify data and research findings. With regard to the second research question, that is to document likely changes of each participant during the study, interview answers and reflective journals were compared either from different participants, or from the same participant at different times in the study (Creswell, 2013).

The interview questions were validated in two ways. Firstly, they were piloted with some undergraduate students before the actual study started (with a different class than the four cohorts which participated in the main study). Secondly, an expert at assessment checked the questions. The principal investigator and the two research assistants analyzed the data. The level of interrater agreement between the main investigator and two researchers was 82% during the coding process, with discussions held in order to reach consensus on final decision. All participants agreed that the quotations as selected and presented in this paper give a fair representation of their experiences and views.

Findings and Discussion

The main findings are presented under the headings of five themes:

Table 2 Main Findings – Main Themes

Theme 1: Grades are more important than feedback because they show what has been achieved. Yet, deciding on grades is more difficult than giving feedback.
Theme 2: Written feedback (in the form of comments) is more important for learning than grades.
Theme 3: Participants' experiences of giving oral and written feedback.
Theme 4: Restricted quality of written feedback.
Theme 5: Feedback related to the assessment criteria. The focus on feedback to show ways for improvement of coursework.

Quotations from the interview answers, reflective journals and the written feedback are provided to illustrate the findings. After each quotation, the reader will find a parenthesis in which there will be one number for each participant (because names have been removed for anonymity purposes), the number of the cohort, the number of the week (1-13), the subject of each participant (Mathematics or Physics) and finally, the data source (interviews or reflective journals). Lastly, any grammar or spelling mistakes in students' quotes are being preserved.

Theme 1: Grades are more important than feedback because they show what has been achieved. Yet, deciding on grades is more difficult than giving feedback.

In the very first and second weeks, the majority of prospective teachers reported that peer-assessment was a unique experience, which they enjoyed. The following statement is representative of such experience:

I think reading and assessing the coursework is exciting and interesting work (8, cohort 1, week 1, Mathematics, reflective journal).

They also explained that peer-assessment, by giving comments and grades, made them confident as they started feeling to be trusted by the teacher and peers. The feeling of being trusted was important and it was further developed during the academic semester. For example:

I have liked being important and being trusted because for all my education years, for much important modules, I've never assessed someone else's coursework. Because in my mind I believed that grading was everything and students or peers could not be trusted, only a teacher can do this (64, cohort 3, week 2, Physics, interview).

In addition, the majority of participants wrote in the reflective journals and explained in the interviews that the provision of feedback is not sufficient. This is because it is mainly grades which inform students about where they are and how they have done. Grades are needed for motivation, too. One representative quote is the following:

Feedback only is not sufficient. Teachers need to give grades because students need to know how they have done. It is grades which give such information (32, cohort 2, week 2, Mathematics, interview).

This is an interesting aspect, which further explains their worry about giving a grade to a piece of coursework. In fact, they considered grading to be more difficult than offering comments. One main reason they gave was that grades are more important for students' success.

At the same time, giving comments was easier because comments do not count for the final grade or for the GPA. Typical statements of such thinking were as follows:

Writing comments was simpler than giving grades in peer-assessment (77, cohort 43, week 1, Mathematics, interview).

The crucial role of grades for one's success made them be careful and reluctant to give grades. One mathematics and one physics teacher, respectively, explained in the interviews:

I feel uncomfortable, and this is the hardest of the two when I should give grades. Giving written feedback like comments is easier for me (81, cohort 4, week 2, Mathematics, interview).

I gave comments to my classmate. However, it was hard to decide what grade to give to my peer. I have changed my decision for a few times (45, cohort 2, week 2, Physics, interview).

Thus, although they enjoyed looking at one peer's piece of work, they wanted to avoid grading. The explanations the participants gave in the interviews help us create a better picture of the limited understanding of feedback and grades and, then, of peer-assessment they have developed in the first weeks of the course. One physics participant stated: *Because you are not allowed to make a mistake. Grades are important for your GPA* (89, cohort 4, week 2, Physics, interview).

Theme 2: Written feedback (in the form of comments) is more important for learning than grades.

During the course, they have practiced peer-assessment whereby they had the chance to offer and receive written comments as feedback and oral feedback, too. Such practice resulted in significant changes in participants' thinking for the majority of them. They got more actively engaged in the process, which made them consider the importance of feedback in the form of written comments for improvement of the coursework. The ideas with which they joined the study changed towards a better understanding of the role of peer-assessment. With regard feedback, many participants started understanding that the issue is not to give grades only but mainly advice about how to improve a piece of coursework.

Giving comments is a much higher skill than giving grades (91, cohort 4, week 3, Physics, reflective journal). And,

I think the most difficult part of the assessment is giving feedback to my peer (54, cohort 3, week 3, Physics, interview).

It has been clear that they changed their views related to grading, as they practiced giving feedback to peer's coursework.

This assessment activity gave us a chance to improve our work. In the case that we only obtained our grades, there would be no improvement. However, feedback from our peers helped us improve our work (69, cohort 3, week 4, Physics, interview).

All these show that there is a clear transition from an initial idea that grades and grading are the most crucial to the idea that written comments are also needed because they provide information about achievement and quality of work. Many times, they wrote in the reflective journals:

Before I got this course, grading was more important, but now feedback (88, cohort 4, week 4, Physics, reflective journal).

Theme 3: Participants' experiences of giving oral and written feedback.

When the participants returned the coursework with written feedback and grades, they had discussions in pairs to discuss written feedback. They reported a range of experiences and benefits from such discussions whereby they shared feedback. The following statement is clear:

We have talked in two different pairs. My classmate asked me a few questions. In my turn, I have asked the points that I did not understand to the person who assessed my piece of coursework (8, cohort 1, week 9, Mathematics, interview).

Such discussions helped them clarify any likely points that were not well understood and make corrections in a supportive environment. Many times they reflected on such discussions they had with peers in pairs. The following statement is representative of the clarification purpose of discussion in pairs:

In this peer-assessment activity, discussions were helpful because we tried to convince each other about our feedback and work. Or, if there are some points that we did not understand well, we had a chance to discuss the point, the topics, as well (51, cohort 3, week 9, Mathematics, interview).

Peer-assessment have led to discussion and maybe disputes between prospective teachers. In such discussions, they learn better and more because they discuss in equal terms. This is because there is no peer who holds any authority (as it is the case with the "traditional" classes in which only the teacher has authority to make decisions). In addition, as classmates,

they communicate in a language that they naturally use. A trusting and a more relaxing relationship has been developed, as they have explained with the following quote:

Discussion with our papers is also helpful because we use the same basic language (78, cohort 4, week 9, Mathematics, interview).

Talking with classmates enables us to re-think about our written work and improve it. We listened to classmates' opinions and why s/he thinks like that. If we get only grades probably we may not be able to improve our work (10, cohort 1, week 5, Mathematics, reflective journal).

Thus, it makes sense that for many of them, such discussions were one of the best parts of the whole process, as the following quotes state:

The best part of the peer-assessment is giving and taking oral feedback from my classmates, communicating with them. I learned from him and also I saw my weaknesses (29, cohort 2, week 5, Mathematics, reflective journal) and,

Talking about peer-assessment with friends is very important because we have the chance to ask: "Why did you give me this grade?" and "How do I improve my work?" and discuss these issues (82, cohort 4, week 5, Mathematics, reflective journal).

All these quotations make clear that participants' discussion in pairs helped them understand the feedback and how they need to revise their coursework. During the semester with practice, they would no longer worry about the peers who were classmates and friends. They enjoyed discussing with them, giving feedback and receiving guidance.

Theme 4: Restricted quality of written feedback.

Under this theme, we have included feedback which was of limited information for the participants. This means written feedback which was restricted to general comments, identification of mistakes and confirmation of right answers. The analysis of written feedback showed that the majority of participants, when started to give feedback, they wanted to confirm correct answers, identify mistakes and give general comments like the following: "*This is a good job*", or confirm right answers like "*I like your answer*" or send a message like "*This answer is wrong"*. In particular, they gave general comments*: Very good. Good work. Nice job. Excellent work. Good effort. Vague. This is a great answer!*

They gave vague recommendations: *More is needed here. Improve this answer. You are not clear here. I do not see what your point is here. And sometimes they gave general advice: Study*

more! This needs improvement. Be more careful. Do your best. Study more. You forgot to give an example. You need more details in your answers.

Theme 5: Feedback related to the assessment criteria. The focus on feedback to show ways for improvement of coursework.

The participants carried on improving their understanding of feedback. In particular, they realized that for feedback to fulfill its specific role to support further learning, it should be related to the assessment criteria. It should, then, identify the strengths and weaknesses (in relation to assessment criteria) and propose ways for improvement. Under this theme we have included types of feedback which are closely related to the assessment criteria and suggestions are explicit. In addition, sometimes such feedback may be put in the form of a question to make the recipients think about how to respond to it. All the following quotations show how feedback can be specific to the subject (Mathematics and Physics) and, to the task; what this study has called subject-specific feedback. The following comments illustrate the high quality subject-specific feedback that prospective teachers gave to their peers (mostly from the mid-term until the end of the academic semester).

Please refer to your graph when you analyze your data.

What about talking about how you will improve your experiment, if you repeat it.

How do you explain your experimental evidence by using your theory?

You forgot to talk about limitations in your experiment.

After you plotted the data points, how have you drawn the curve?

What do you think about the forces acting upon the mass? Can you sketch them?

I like the way you organized your ideas and examples in your answer.

This is a complete answer because you have included all the possibilities of different line graphs.

Think about what the learning targets are for today's lesson.

When reflecting on the process of feedback they wrote:

Peer assessment is a good way to let your classmates know about their weaknesses and strengths because we are in the same class (70, cohort 3, week 5, Physics, reflective journal).

I was very pleased while I am/was examining my peer's paper. Actually, peer-assessment is very beneficial because when we write our papers, we cannot see where we have any

drawbacks but when we assess our friend's work, we can easily see the strengths and weaknesses of her paper (54, cohort 3, week 5, Mathematics, reflective journal).

The above quotations provide much information about the process and its benefits to both parts (assessors and assesses). In fact, it is easier for anyone to identify mistakes and weak points in the peer's work than in his own and, then, come back to his/her own paper and make the relevant corrections. At the same time, the assessor, when looks at a better piece of work, can take ideas from strong points and improve his/ her own work. Also, we need to underline that the participants (by time and practice) have supported each other, they have collaborated in order to learn better and improve their work. There are no longer worries about how not to offend classmates and so on. They have come through a range of learning experiences in which they developed confidence that together they can learn better. They take time to use feedback and improve their work as the next quotations show:

With my peer's comments, I will improve my answer and learn things that I do not know (34, cohort 2, Mathematics, interview).

Peer-assessment has given new ideas to me. I looked at my own answers from my classmate's different perspective (39, cohort 2, Mathematics, interviews).

Thus, peer-assessment is a two-way process. Each participant has helped his/her peer with the written feedback they offer and at the same time, they get ideas about how to improve their own coursework. Finally, they underlined the learning aspect of the process and their attention to the assessment criteria:

What else is needed here so that you explain your ideas better? (Hint: think about the assessment criteria) (written feedback)

Since we learned about the assessment criteria in the lesson, applying them when looking at the work of our classmates was a great learning experience (11, cohort 1, week 12, Physics, reflective journal).

"*I learn better by giving feedback*" (47, cohort 2, week 12, Physics, reflective journal).

This is very important. This is true because when they give feedback, it is easier for them to identify mistakes in the peer's work more easily than in their own. Many prospective teachers wrote:

I found my friend's coursework better than mine. I learned from his paper (52, cohort 3, week 12, Physics, reflective journal).

The real learning happened when I evaluate the strengths and weaknesses of another person [...] *When you realize your strengths and weaknesses, you can improve your work in a wonderful manner* (77, cohort 4, week 12, Mathematics, reflective journal).

The meaning of "wonderful" is to do with immediate and explicit ideas for improvement and the fact that they have collaborated with their classmates with whom they speak the same language. They have talked explicitly about the process of understanding the assessment criteria better and applying them. The participants discuss the assessment criteria and then they assess peer's coursework. The process of giving subject-specific feedback and discussing in pairs, resulted in participants' understanding and learning the assessment criteria and what is required for a high quality work.

For most of them the process was interesting since they would need to think about the expected quality of answers and focus on the related assessment criteria.

… applying the assessment criteria to assess our classmates' work was a great learning method…. (89, cohort 4, week 12, Physics, interview) and,

Since we discussed the assessment criteria of our written work, applying them to assess our friends' work was a very helpful experience. Now I much better understand how some work is assessed. The process was interesting (27, cohort 2, week 12, Mathematics, reflective journal)

Having gone through the practice of the peer-assessment process, they concluded*:*

Peer-assessment is an intense and time-demanding process. It requires, firstly, good understanding of the topic to enable you to assess effectively. Secondly, it requires time to go through and assess the work and time to give constructive written feedback (7, cohort 1, week 13, Mathematics, interview).

By participating in peer-assessment, I learned many things: how to assess, how to give feedback and how to discuss our feedback with our peers (52, cohort 3, week 13, Mathematics, reflective journal).

Conclusions and Suggestions

The practice of peer-assessment has opened the way to a wide range of experiences for prospective mathematics and physics teachers. As they wrote in the reflective journals and explained in the interviews, they did like the process of giving comments and grades. It was the first time they were given such an opportunity; this was, so far, solely the instructor's duty. They experienced the feeling of being powerful to give grades and being trusted by the teacher and

classmates. First of all, the findings, as already presented in the previous section, agree with those by similar studies on assessment in Turkey (i.e. Gelbal & Kelecioğlu, 2007; Koç, 2011; Özdemir, 2016).

One main finding of the present study is that in the first weeks of the course, prospective teachers demonstrated a limited understanding of peer-assessment due to a limited understanding of comments and grades. The written comments they started giving were restricted to general ones: "*This is a good job*", or to confirm right answers like "*I like your answer"* or to identify wrong answers "*This answer is wrong*". They also supported the view that grades are more important for students, because grades do matter for success. In addition, the participants supported the view that grades offer better information about student achievement than feedback does. Thus, the majority of them appeared to give simple comments and also not to feel comfortable to assign grades. For them, writing simple comments does not require much time and effort; it is easy and simple. In addition, according to them, giving grades requires more thinking and effort. Taking it as a whole, their understanding of peer-assessment seemed to be limited as it is clear that they did not realize the full potential of feedback; for how to improve a particular piece of coursework and for how to support current and future learning. This very first finding of the present study seems to be in contradiction with what research has shown. In the literature review section, it was made clear that Butler (1988), among more researchers, provided strong evidence that marks and feedback (when provided together) do not help students improve their work. This is because they get distracted by marks and thus, they do not pay any attention to the comments offered to them for improvement. Or, grades only, do not contribute to learning because they do not give them any information about current achievement or further learning. However, studies like the one carried out by Smith and Gorard (2005) reported on an intervention in which they did not give any grades to students but only comments to help them understand what they needed to improve. Although, according to the intervention, students should have focused their attention only on comments, the participants claimed their marks because it was marks that they could understand better.

On the other hand, it seems that there is one paradox here. The participants' idea that giving feedback is more straightforward than giving marks reflects a naive idea, which ignores the crucial role of feedback in the learning process. "Writing comments was simpler than giving grades in peer-assessment". While prospective teachers were not familiar with feedback in the form of comments, but with grades only, they stated that giving feedback was easier than giving grades. They seemed to hold a naive idea of comments, not related to the role of feedback in

support of learning. Interviews shed more light to such experiences and perceptions so that we better understand their reasoning. Such an experience may be interpreted by considering that until that time, their experiences were only related to summative assessment and grades. Indeed, they joined the course with "initial" ideas that marks are important for one's success and for the total GPA. The dominance of summative assessment resulted in a limited understanding of peer-assessment because their experiences originated from grade-dominated teaching. Secondly, peer-assessment was carried out in the class environment with 22-26 prospective teachers with classmates as peers. This resulted in them not feeling comfortable to do peer-assessment by giving grades to classmates with whom they were friends.

Through the development of the course, many of the participants started experiencing that grades by themselves cannot give any information about further learning. When they realized what the meaning and the role of feedback are for the learning process, they emphasized that giving feedback is more important. Many of the participants were concerned about good quality feedback and not concerned about grades. They started recognizing the value of the peer-assessment approach. They recognized that feedback should not focus on the person and this delicate notion helped them proceed when the peer was their classmate. This makes sense because the classes, in which the study took place, were not large. The development of an appropriate understanding of the role of feedback would take a whole academic semester. In addition, the present study provided strong evidence that there was a development in the type of feedback they gave to each other. From the third week onwards, their experiences shifted towards better peer-assessment practice. In parallel with the progress of the assessment course, they experienced that they can have a role in the assessment process; this idea could not be shared by lecturing only. They started developing a better understanding of the role of feedback in the learning process. They started adopting research-based assessment strategies; mainly that feedback is more important for one's learning than grading.

It took them enough time to start thinking in terms of the assessment criteria and the required quality so that written feedback to be related to the assessment criteria. It required also much effort for them to be convinced that firstly, they need to think about the provision of good feedback and then, about grades in support of learning (Brookhart, 2017). Their understanding of feedback improved when they started experiencing that feedback (in relation to assessment criteria and the required quality) is more crucial for one's learning. Thus, feedback should not be restricted to show right or wrong answers but to identify strong points, weaknesses and missing points. Feedback should then show ways for improvement. In other words, feedback

should enhance learning by highlighting the strengths and weaknesses and also by giving suggestions about what needs to be done so that coursework will get improved.

Discussions with peers in pairs helped them advance their understanding of assessment criteria and subsequently, of feedback. Discussions seemed to be an intermediate "step" between simple types of feedback (Theme 4) and feedback related to the assessment criteria (Theme 5). The participants explained how the discussions in pairs helped them improve coursework and learn better. This was because the discussions had a real purpose: for them to have the opportunity to communicate written feedback, to clarify likely difficult points and suggest ways for improvement. Prospective teachers considered it as the most valuable part of the peer-assessment process, because they felt comfortable to talk to their peers about their coursework. In fact, they were much more comfortable than talking to the teacher. They would ask questions to their peer more easily than to the teacher. They take responsibility for their own learning on an equal basis with their peer. The participants in this study enjoyed discussing on the basis of their work as they would talk and use the same natural language. They see that they can take some responsibilities and make decisions about their peer's learning, which, so far, was the responsibility of the teacher only. They have worked within the zone of proximal development to "scaffold" peer's learning (Vygotsky, 1986). Through peer interaction, successful peer feedback is being communicated to support the learning process. In discussions, prospective teachers attend to each other as resources for learning (Wiliam & Leahy, 2015). The process becomes transparent because they are able to understand where they are in their learning, where they need to go and how best to get there (Sadler, 1989). Learning how to give and accept comments, how to make suggestions, how to justify one's position and reject suggestions are all important skills. A full understanding of feedback is achieved when the aim for each of them is to advance the quality of peer's coursework. They take time to use feedback and improve their work. Such an experience is important because the theory on classroom assessment highlighted the need for the feedback to be understood by students and then, to be used for improvement of coursework and for a second submission (Black & Wiliam, 1998). Their engagement in applying the assessment criteria resulted in them giving subject-specific feedback, which by itself, is important knowledge for prospective mathematics and physics teachers.

Notably, there is one more important issue related to the practice of giving written feedback. Feedback seems to be immediate; in a short time after coursework submission. The participants would not need to wait for a certain time as it is the case when feedback is provided by the teacher. Teacher feedback is mostly not offered in time for students to benefit

immediately. Right after discussions, they would get time to work on it and improve their work. The whole process, although it may seem to be time-consuming, it is not. Much time was saved. But in peer assessment feedback is given immediately by peers so that they are able to act upon it and improve their own work.

The progress that prospective teachers made in giving feedback (from general comments to subject-specific comments) made them more critically aware of the process of peer-assessment. The participants experienced that comments guiding the learner toward the next step are more valuable, as they promote learning. Consequently, the idea and confidence that they can improve attainment is apparent. All of them participated to do their best. They were actively involved in applying the assessment criteria and demonstrating the expected quality they would need to demonstrate in the coursework. Thus, they explained that this sort of peer-assessment helps them shift the emphasis towards improvement of coursework. They developed the idea and confidence in themselves that they can improve attainment, and then how to help their peer.

Giving feedback and grades requires, first of all, good subject knowledge (they may need to go back to their notes) and development of assessment knowledge. Consequently, they gained confidence in providing feedback in their own subjects, Mathematics and Physics. In fact, they gave good quality feedback to their peers; both written and oral (when discussing in pairs). This experience has been rewarding: to write comments on strengths and weaknesses and give ideas about how to improve their peers' and their own coursework. The whole process is double-sided because both parts have benefited: the assessor and the assessed peer. The emphasis was not only on the final product (the final version of coursework) but primarily on the process. They also learned that good quality feedback is specific to the subject (Mathematics and Physics), the specific tasks and coursework.

The present study provided a range of challenging experiences with regard classroom assessment in initial teacher education. It also provided strong evidence that the course improved their assessment knowledge and assessment skills to some extent. On the basis of such research findings, we want to argue that the development of assessment knowledge and research-based assessment skills does take much time. In fact, more practice is needed as there was not enough time for them to develop skills towards the direction of grading. The duration of 13 weeks was not enough. A second assessment course is needed. In fact, more time is needed for the prospective teachers of the present study so that they develop an adequate understanding of grading in support of learning.

This study highlights the importance of including assessment education and development of assessment knowledge and skills in initial teacher education. Shepard and her colleagues (2005) strongly argued that prospective teachers should be taught about assessment which supports learning by practicing certain research-based assessment methods and giving feedback to their own work. The study, by its design, did not attempt to generalize findings to all prospective teachers. Such an aim would be beyond the context of the present study. We should also consider the constraints of one academic semester. We would like to carry on the analysis of the data in order to compare the participants' comments and grades with those of the instructor's.

Understanding of prospective teachers' experiences, difficulties and challenges during their practice is necessary for the design of effective initial teacher education programs. Prospective teachers experience difficulties and have needs which must be considered when planning and implementing initial teacher education. Prospective teachers need help to move from the focus on grades only, towards the learning of assessment criteria and finally, to focus on written feedback - comments related to assessment criteria. Prospective mathematics and physics teachers need support to apply assessment criteria to the specific subjects of Physics and Mathematics. The role of the Departments of Mathematics and Physics Teaching in developing assessment practices in prospective teachers is very important.

Further research is needed to follow the same participants when they obtain teaching posts in schools to identify the range of needs they experience in the first years of teaching in schools. Another suggestion would be to continue working with the same participants to further develop classroom assessment skills. Last, but not least, we would like to invite Mathematics and Physics educators to investigate similar classroom assessment issues with prospective mathematics and physics teachers at different universities.

Acknowledgements

Special thanks to all prospective teachers who participated in the study and gave much effort and time from their busy schedule. To the two research assistants who helped with data analysis.

Appendix A

Main Topics of the Assessment Course

Classroom Assessment

Development of learning targets

Development of assessment goals and criteria

Development of assessment tasks

Validity and reliability of assessment

Alignment among curriculum, instruction and assessment

What makes feedback effective?

Feedback and Grades: What does research say?

Assessment for Learning

Grading- Using learning targets to guide summative assessment and grading

Performance Assessment

Appendix B

Questions in semi-structured interviews

What is your experience with peer-assessment?

Did you like it?

Why? Yes? Or why not?

What is it that you liked?

What is it that you did not like?

What did you learn?

Did you help your peer? How?

Did your peer(s) help you? How?

What was difficult for you?

Was giving feedback difficult? Why?

Was giving comments difficult? Why?

How did it help you? Was giving grades difficult?

What did you learn? Please give me an example.

Are you going to use peer-assessment when you are teacher in a school?

What sort of classroom assessment have you found important?

How confident are you in developing assessment criteria?

How did feedback help you find out how to do better? How did peer-assessment help you?

Matematik ve Fizik Öğretmen Adaylarının Akran Değerlendirmesini Uygulama Deneyimleri

Özet: Sunulan araştırmanın amacı, matematik ve fizik öğretmen adaylarının akran değerlendirmesini nasıl deneyimlediklerini araştırmaktır. Çalışma, hizmet öncesi öğretmen eğitimi için tasarlanmış bir sınıf içi ölçme-değerlendirme dersi kapsamında gerçekleştirildi. Çalışmaya dört gruptan doksan dört (94) öğretmen adayı katıldı. Katılımcılar akranlarının ödevlerine geri bildirim verdiler, yansıtıcı günlükler yazdılar ve yarı yapılandırılmış görüşmelere katıldılar. Katılımcılar, not vermenin geribildirim vermekten daha zor olduğunu, çünkü notların öğrencilerin başarısı için "sayıldığı" için önemli olduğunu bildirdi. İkinci olarak, ders boyunca katılımcıların düşünceleri, yorum vermenin not vermekten daha önemli olduğu fikrine "kaydı" çünkü geribildirim onların değerlendirme kriterlerini anlamalarına ve ardından dersleri iyileştirmelerine yardımcı oldu. Bulgulara dayalı olarak, makalenin ana argümanı, öğretmenlerde sınıf içi değerlendirme becerilerinin ve değerlendirme bilgilerinin geliştirilmesinde hizmet öncesi öğretmen eğitiminin rolü hakkındadır.

Anahtar kelimeler: değerlendirme kriterleri, geri bildirim, not verme, akran değerlendirmesi, hizmet öncesi matematik ve fizik öğretmen eğitimi.

References

Acar-Erdol, T., & Yıldızlı, H. (2018). Classroom Assessment Practices of Teachers in Turkey. *International Journal of Instruction, 11*(3), 587-602. https://doi.org/10.12973/iji.2018.11340a

Black, P. (2003). *Testing: Friend or Foe? Theory and Practice of Assessment and Testing*. Routledge Falmer Press.

Black, P. & Harrison, C. (2004). *Science inside the black box: Assessment for learning in the science classroom*. King's College and GLAssessment.

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80* (2), 139-144.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education, 22*, 151-167. https://doi.org/10.1080/713695728.

Boud, D. (2001). Using journal writing to enhance reflective practice. *New Directions for Adult and Continuing Education, 90*, 9-17.

Boztunç Öztürk, N., Şahin, M.G., & Ilhan, M., (2019). An analysis of scoring via analytic rubric and general impression in peer assessment. *Turkish Journal of Education, 8*(4), 258−275. DOI: 10.19128/turje.609073

Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice, 30*(1), 3-12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Brookhart, S. M. (2017). *How to Use Grading to Improve Learning.* ASCD Publications.

Brookhart, S. M & McMillan, J. H. (2020). Introduction. In S. M. Brookhart and J. H. McMillan (eds.) *Classroom Assessment and Educational Measurement* (pp. 1-7). Taylor and Francis Publications.

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T. & Welsh, M. E. (2016). A Century of Grading Research: Meaning and value in the most common educational measure. *Review of Educational Research, 86*(4), 803-848. https://doi.org/10.3102/0034654316672069.

Butler, R. (1988). Enhancing and undermining intrinsic motivation; the effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology, 58* (1), 1-14. https://doi.org/10.1111/j.2044-8279.1988.tb00874.x

Campbell, C. & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues and Practice, 26*(1), 9-18. https://doi.org/10.1111/j.1745-3992.2007.00084.x

Creswell, J. W. C. (2013). *Qualitative inquiry and research design. Choosing among five approaches* (3rd ed.). Sage.

Creswell, J. S. & Miller, D. L. (2000). Determining validity in qualitative inquiry. Theory into Practice, 39 (3), 124-130. https://doi.org/10.1207/s15430421tip3903_2

DeLuca, C. (2012). Preparing Teachers for the Age of Accountability: Toward a Framework for Assessment Education. *Action in Teacher Education, 34*(5-6), 576-591, DOI: 10.1080/01626620.2012.730347

Falchikov, N. & Goldfinch, J. (2000). Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70*(3), 287-322.

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

Hattie, J. (2012). *Visible learning. Maximizing Impact on Learning*. Routledge. Taylor and Francis Publications Group.

Hodgen, J. & Wiliam, D. (2006). *Mathematics inside the Black Box: Assessment for Learning in the Mathematics Classroom*. GLassessment.

Kayacan, A., & Razı, S. (2017). Digital self-review and anonymous peer feedback in Turkish high school EFL writing. *Journal of Language and Linguistic Studies, 13*(2), 561-577.

Ketonen, L., Hähkiöniemi, M., Nieminen, P. & Viiri, J. (2020). Pathways through Peer Assessment: Implementing Peer Assessment in a Lower Secondary Physics Classroom. *International Journal of Science and Mathematics Education, 18,* 1465–1484 https://doi.org/10.1007/s10763-019-10030-3.

Koç, C. (2011). The Views of Prospective Class Teachers about Peer Assessment in Teaching Practice. *Educational Sciences: Theory and Practice, 11* (4), 1979-1989. ERIC Number: EJ962684

Kvale, S. & Brinkmann, S. (2009). *InterViews : Learning the craft of qualitative research interviewing*. Sage Publications.

O'Donnel, A. M. & Topping, K. J. (1998). Peers assessing peers. Possibilities and problems. In K. J. Topping and S. Ebly (eds.). *Peer-assisted learning* (pp. 255-278). Lawrence Erlbaum Associates.

Ozmen, K. S., & Aydın, H. U. (2015). Examining student teachers' beliefs about oral corrective feedback: Insights from a teacher education program in Turkey. *Australian Journal of Teacher Education (Online), 40*(12), 141–164. doi/10.3316/informit.681001963960425

Özdemir, S. (2016). The opinions of prospective teachers on peer assessment. *Educational Research and Reviews, 11*(20), pp. 1859-1870. DOI: 10.5897/ERR2016.2997

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18,* 119-144. https://doi.org/10.1007/BF00117714.

Sadler, R. (1998). Formative Assessment: revisiting the territory. *Assessment in Education, 5*(1), 77-84. https://doi.org/10.1080/0969595980050104.

Sasmaz Oren, F. (2012). The effects of gender and previous experience on the approach of self and peer assessment: a case from Turkey. *Innovations in Education and Teaching International, 49* (2), 123-133, DOI: 10.1080/14703297.2012.677598.

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

Shepard, L., Hammerness, K., Darling-Hammond, L., Rust, F. with Snowden, J. B., Gordon, E., Gutierrez, C. & Pacheco, A. (2005). Assessment. In L. Darling-Hammond and J. Bransford, J. (eds.) *Preparing Teachers for a Changing World. What teachers should learn and be able to do* (pp. 275-326). San Francisco, CA: Jossey-Bass Wiley National Academy of Education.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78* (1), 153-189. https://doi.org/10.3102/0034654307313795.

Silverman, D. (2001). *Interpreting qualitative data* (2nd edition). Sage.

Sluijsmans, D. M. A. & Prins, F. (2006). A conceptual framework for integrating peer assessment in teacher education. *Studies in Educational Evaluation, 32* (2006), 6-22. https://doi.org/10.1016/j.stueduc.2006.01.005

Sluijsmans, D. M. A., Brand-Gruwel, S. & van Merriënboer, J. J. G. (2002) Peer Assessment Training in Teacher Education: Effects on performance and perceptions. *Assessment and Evaluation in Higher Education, 27*(5), 443-454. DOI: 10.1080/0260293022000009311.

Smith, E. & Gorard, S. (2005). 'They don't give us our marks': The role of formative feedback in student progress. *Assessment in Education: Principles, Policy and Practice, 12*(1), 21-38, DOI: 10.1080/0969594042000333896

Stake, R. E. (1995). *The art of case study research*. Sage.

Şahin, S. (2008). An application of peer assessment in higher education. *The Turkish Online Journal of Educational Technology, 7*(2), 5−10. ERIC Number: EJ1102895

Şahin, M. G., Taşdelen Teker, G. & Güler, N. (2016). An analysis of peer assessment through many facet Rasch model. *Journal of Education and Practice, 7*(32), 172−181. ERIC Number: EJ1122446

Topping, K. J. (1998). Peer assessment between students in colleges and university. *Review of Educational Research 68*(3), 249-276. https://doi.org/10.3102/00346543068003249

Topping, K. J. (2009). Peer assessment. *Theory into Practice, 48,* 20-27 https://doi.org/10.1080/00405840802577569

Tseng, S-C. & Tsai, C-C. (2007). On-line peer assessment and the role of the peer feedback: a study of high school computer course. *Computers and Education, 49*(4), 1161-1174. https://doi.org/10.1016/j.compedu.2006.01.007

Vygotsky, L. S. (1986). *Thought and Language*. MIT Press.

Wiliam, D. & Leahy, S. (2015). *Embedding Formative assessment: Practical Techniques for K-12 Classrooms.* LearningSciencesInternational.

Yin, R. K. (2007). *Case study research and applications. Design and methods*. Sage.