

Remote Assessment in Higher Education during COVID-19 Pandemic

Selma Senel ^{1,*}, Huseyin Can Senel ²

¹Balikesir University, Faculty of Education, Department of Educational Sciences, Balikesir, Turkey

²National Defense University, Army NCO Vocational College, Department of Computer Technology, Balikesir, Turkey

ARTICLE HISTORY

Received: Nov. 02, 2020

Revised: Dec. 31, 2020

Accepted: Jan. 30, 2021

Keywords:

Remote assessment,
Measurement and
evaluation,
Distance education,
Online test,
E-assessment.

Abstract: Universities have made a compulsory shift to distance education due to the Covid-19 pandemic. All of the higher education institutions in Turkey have completed 2019-2020 Spring semester using online tools. However, most of these institutions were not fully-prepared to have all of their courses online. Technical inadequacies, lack of qualified online tools, inexperience of instructors and students in distance education have emerged as major issues that institutions have to face. In addition to all, a new question arised; which approaches will be used for assessment. This study aimed to seek the common assessment approaches used through pandemic, how students perceived the quality of the assessment and the pros and cons of using these practices. Additionally, we examined whether participants' perceptions about quality of the assessment differ according to interaction with faculty members and use of online tests. Researchers employed survey design to reply four research questions and used a three-part instrument to collect qualitative and quantitative data. 486 students from 61 universities voluntarily participated in the study. Results indicated assignments are the mostly used tools and students are generally satisfied about the quality of the assessment practices. Another result is that students who interact with faculty members are more satisfied with the quality of the assessment practices. This emphasizes the importance of formative assessment and feedback in remote assessment. Further, students who took online tests are more satisfied with the quality of assessment. Suggestions were made for future research.

1. INTRODUCTION

Throughout history, pandemics are known to affect human life in many ways (Martini et al., 2019). The COVID-19 pandemic, which we still largely feel, has also caused critical changes and it also has engendered significant transformation in education activities all over the world (Daniel, 2020). Countries where the COVID-19 pandemic threat has increased, conventional education have been suspended temporarily and the distance education tools were adopted (Bozkurt & Sharma, 2020). In Turkey, as in primary and secondary education institutions affiliated to the Ministry of National Education, higher education institutions have completed 2019-2020 spring semester using distance education. A similar decision was taken for 2020-2021 fall semester.

CONTACT: Selma ŞENEL ✉ selmahocuk@gmail.com 📍 Balikesir University, Faculty of Education, Department of Educational Sciences, Balikesir, Turkey

ISSN-e: 2148-7456 /© IJATE 2021

It is impossible for 21st century educational institutions to use a method away from technology. Today, innovative tools are commonly used both in-class and out of-class activities (Akçayır & Akçayır, 2018). Distance education is more common thanks to these tools and the number of distance education institutions is increasing. However, a compulsory transition to distance education without adequate preparation may cause problems in different aspects of distance education. Providing the necessary technical infrastructure for distance education, utilizing technological tools and having experienced teaching staff in sufficient numbers in distance education is among the basic needs of distance education (Veletsianos & Houlden, 2019). Absence of basic needs can be predicted to negatively affect the quality of distance education and the extent of this effect is worth researching.

Valid and reliable assessment results are crucial to be able to control whether the educational goals have been achieved or not. Assessment can be carried out during the training in order to identify and then eliminate learning deficiencies as formative assessment. The instructor may explain the assessment results and give feedback. In this respect, assessment practices have important effect in the achievement of educational goals (Chen et al., 2020). In addition, summative assessment have guiding impact by forming a basis for decisions such as being successful in a course, moving to a higher education institution, receiving a diploma or certificate (Biesta, 2009). With these in mind, both formative and summative assessment practices are considered as the cornerstones of instruction.

Formative assessment can be expected to be more prominent in distance education since the students are 'remote' and the possibility of interaction is low. Since there is no conventional classroom environment, the student needs feedback in order to see their deficiencies and mistakes. This requires effective interaction between student and instructor. Instructors should be able to provide students with the opportunity to organize their learning by providing instant feedback, through tests or performance-based techniques (Hatzipanagos & Warburton, 2009). To summarize, "monitoring" and "feedback", which is a part of formative assessment in distance education is gaining more importance. Feedback can be considered as the primary means of student-faculty communication and interaction.

All of the universities in Turkey have completed 2019-2020 spring semester with online tools. Assessment practices were conducted using various techniques like online tests, assignments, and projects. There was no face to face exams. In this period, a new issue has arisen about the quality of the assessment carried out with online tools. Assessment results form students' grade point averages and graduation besides the general achievement goals. In other words, the critical decisions that may affect the lives of individuals were made based on the assessment results and it was the first time that all assessment practices were made upon distance tools.

Learning management systems are widely used in distance education. These tools provide integrated functions like communication, interaction and storage. Canvas, Blackboard, Edmodo, Moodle, Google Classroom and Microsoft Teams are some of these tools. Similarly, video conference tools like Zoom, Skype and Adobe Connect (Koh & Kan, 2020; Nyachwaya, 2020) is latest tools that are common to have online lessons. In addition, these tools can provide a number of advantages for assessment (Araka et al., 2020). The advantages of using these tools in assessment are listed as follows:

Instant feedback: It is known that using instant feedback increases the performance of the students in summative assessment (Joint Information Systems Committee, 2010; Shrago & Smith, 2006). Therefore, feedback on assessment results has a critical role in increasing the quality of the learning. Among the tools used in distance education, tests using items that require selection (multiple choice, true-false, matching), are very appropriate for producing instant feedback. Using instant feedback, students may find opportunity to organize self-learning by noticing deficiencies and mistakes.

Ease of editing based on feedback: An assignment submitted electronically is easy to examine and edit. Students can comfortably edit and re-organize assignments in line with instructor's feedback. Instructor may plan the re-submission of assignments and students may re-submit the latest version of their work.

Ease of submitting/responding: Most of the learning management systems have testing or delivery tools which response and product delivery can be systematically and easily carried out. These tools are widely used for remote-assessment (Moore et al., 2011). In addition, common technological tools such as e-mail or direct messages also offers delivery preferences. Uploading or submitting an assignment to a web-based tool is easier and faster for students to maintain and submit the physically formed product.

Control and storage: Online storage, access, and control of tests and assignments are easier with distance education tools. Informative data such as the list of the submitted/missing assignments, submission date and time are automatically kept in most of the distance education tools. The faculty member may save all test documents to internal storage devices (computer, portable disk, etc.) or reach them independently from time and place.

Providing statistical data: Besides providing test statistics, distance tools present data about students' participation rates. Although there is important debate about the relation between access rates to learning management systems and completion of course outcomes, instructors may use access or participation data such as access rate, participation time, message rate and message length to gain insight (Murray et al., 2012).

Potential to enrich assessment tools and products: The ability to use media such as images, graphics, drawings, audios, videos and animations provided by latest technology can provide richness in assessment by changing assignment framework (Williams et al., 2005). The instructor may submit an animation and ask students to prepare a video as a reflection assignment and share this video on social media to raise the awareness of the society on related subject.

Providing student participation and motivation: Computer-based assessment practices, which are able to use interactive techniques and include multimedia such as audios, images, animations and videos may help to increase students' motivation (Cheng & Basu, 2006). In addition, it is well-known that use of instant feedback increases student participation and motivation in distance education (Chaiyo & Nokham, 2017).

Re-use: It is simple to copy or re-use an online test or assignment prepared with online tools. As reported above, storage and access to data are limitless and instructors may safely share assessment tools with each other.

In addition to the advantages of use of online tools in assessment, there are also some limitations. The most controversial topic in remote assessment is *test security* (Rovai, 2000). Test security is a critical issue to be able to rely on the test results. Test security is exceptionally important when results are used for critical decisions such as student selection, placement and graduation due to the fact that these decisions have high impact and accountability (Frey, 2018). Preventing cheating, copying and plagiarism in assessment in distance education is challenging. This may overshadow the fairness and reliability of the results obtained by assessment. To prevent this problematic situation, different technologies such as voice and retinal scans have been developed (Jain et al., 2006). However, these high-tech solutions have not yet become widespread. On the other hand, in order to prevent cheating and to increase test security, some other handy techniques may be used such as adding time limitations in online tests, presenting test items or choices randomly (in different order), creating an item pool and presenting random questions to each student, making exams using an open camera (proctoring) and hindering new

web pages/tab (Arnold, 2016; Peterson, 2019). However, *complete test security* is not yet possible even all of these measures are provided.

In addition to new technologies and techniques stated above, preferring appropriate assessment techniques and tools may be another option for higher test security (Nguyen et al., 2020). Some of these techniques may be aligned as assignments, take-home exams, performance tasks, e-portfolios and peer/self assessment forms. However, these tools must be activating higher level skills. In other words, these tools must include items or tasks triggering student's thinking, criticizing, evaluating, creating an idea or product, while preparing students for related tasks or questions. Items and tasks must be unique and must create possibilities to reply with autonomous effort. Otherwise, students may copy from web or from other sources (Rowe, 2004). Rubrics, rating scales and control lists may be used for scoring these tools. Using take-home and open-book exams (Atilgan et al., 2009) is another alternative tool. Open-book exams which allow utilizing books, notebooks and other materials may help to decrease cheating. Take-home exams may be considered as a good example for open book exams.

Besides the security limitations, ICT literacy is another competence for assessment in distance education. The ability of faculty members and students to use technology and related tools or the limitations of these devices (computers, mobile tools, internet) may adversely affect the qualified utilization of assessment tools. Participants should have all the technical infrastructure like software, hardware, and internet connection. Problems in connection speed, disconnection or other technical problems can cause hard-to-compensate results, especially in online tests. Performance-based approaches, which are time independent, can reduce the negative effects of technical deficiencies.

It cannot be denied that computer technologies have created informative, facilitating, and accelerating advantages for developing or using online assessment techniques. However, it should be kept in mind that it is up to faculty members to develop valid and reliable measurement and evaluation. Developing a valid and reliable test is incomparably important to which technology is used. Test designers must consider validity and reliability of the test rather than the type of the online tool.

During the pandemic, faculty members necessarily carried out distance education for all courses and all of the assessment practices were conducted online. However, they had been experienced in face-to-face instruction and they are not fully experienced in neither distance education nor remote assessment. Inexperience, technical problems, or lack of expert personnel might have adversely effect distance education period. Some other limitations may have negative effects on distance education and particularly on assessment. For example, the limitations of the learning management systems or decisions of the administration might have hindered preferences of faculty members. For these reasons, reliability and validity of the assessment results might have been in differentiated. Providing a shot about assessment practices carried out in this very first phase of pandemic will be an important indicator for results of remote assessment and will shed light for the future applications.

The purpose of this research is to examine the assessment practices of universities during the Covid-19 pandemic. For this purpose, answers will be sought for four research questions:

1. How are the higher education students' perceptions about the quality of assessment practices carried out during the Covid-19?
2. What are the assessment approaches that higher education institutions prefer during the Covid-19?
3. What are the views of higher education students about the assessment practices carried out during the Covid-19?

4. Do participants' perceptions about the quality of assessment practices differ according to the interaction with faculty members and use of online tests?

2. METHOD

The main aim of the research described in this paper was to present the assessment practices used by universities during COVID-19 pandemic and how students experienced this unique period. This includes gaining an understanding of what practices (distance tools) universities used for assessment, how they used these tools and then to determine the views of students about assessment practices. A survey model was employed in this research using quantitative and qualitative data together. The data obtained for this study consists of the responses from 486 participants from 61 different universities who took distance education for a semester and were evaluated using distance tools.

2.1. Study Goup

The study group for this research was determined through convenient sampling. Undergraduate students who are studying at different universities were reached through the social circle of researchers and social networks. They were informed about the research and volunteering students were identified. The study group, consists of 486 students from 61 universities and 69 departments. Since there were too many universities and reporting the names of all 61 universities would not be a necessary and useful data for the research, universities were grouped considering the University Ranking by Academic Performance (University Ranking by Academic Performance [URAP], 2020). Therefore, "university rankings", which rank universities according to various criteria, were used in reporting the universities participating in the study. The universities participating in this study were analyzed according to 11 different "university rankings list" (URAP Turkey, 2020). Being listed in "university rankings list" can provide information about the quality of universities. Accordingly, it was seen that some of the universities participated in this study were not included in any of the "rankings", while some were included in all of the 11 "rankings". Table 1 summarizes the rankings of the universities that participated in this study.

Table 1. *Distrubition of the universities and faculties according to "university rankings".*

Faculty	0-2	3	4-8	9-11	Total	Percent
Education	10	64	37	77	188	38.68
Arts and Science	10	3	9	3	25	5.14
Fine Arts	3	0	1	1	5	1.03
Law	1	1	2	1	5	1.03
Economics and Administrative Sciences	10	8	4	9	31	6.38
Engineering	8	74	27	8	117	24.07
Medicine	4	7	4	14	29	5.97
Tourism	0	73	1	12	86	17.70
Total	46	230	85	125	486	100.00
Percent	9.47	47.33	17.49	25.72	100.00	

As can be seen in Table 1, study group consists of the students from 8 different faculties. 25% of the participants study at universities which are ranked 9-11 in the university ranking lists. More than half of the participants study at universities which are ranked 0-3 of the university ranking lists. This can indicate that a study group studying at universities with different qualifications. Gender and grades of the participants were summarized in Table 2.

Table 2. Gender and grade distribution of study group.

Grade	Female	Male	Total	Percentage
1	74	38	112	23.05
2	64	42	106	21.81
3	111	48	159	32.72
4	61	36	97	19.96
5	5	1	6	1.23
6	3	3	6	1.23
Total	318	168	486	100.00
Percentage	65.43	34.57	100.00	

Table 2 reveals that that the study group is predominantly composed of female (65.43%). In addition, the frequency of 5th and 6th grades are low. This stems from that undergraduate programs are mainly 4 years in Turkey.

2.2. Data Collection

The data collection tool used in this study is developed as a single form. However, data collection tool includes three main parts. The aims and properties of each part of the tool is explained below.

Part I. In the first part of the data collection tool, an 11-item instrument was used to determine the students' perception about the quality of the assessment practices carried out during pandemic. As the first step of the scale development procedure, literature about the assessment in distance education were reviewed. Then, first version of the items was written considering the basic principles to be followed in the measurement process of a course. A total of 11 items were written. Instrument is a 5-point Likert type ranging from (1) *totally disagree*, (2) *disagree*, (3) *partially agree*, (4) *agree*, (5) *strongly agree* response categories.

Since the data collection tool is applied as a single form; expert views and pre-trial applications were carried out together for all parts of the tool. The views of three experts from *measurement and evaluation in education* department and two experts who have studies in distance education were consulted and improvements were made in the form. Data about universities, faculties, departments and gender, grade, grade point averages were added to the form to be able to describe participants. The form was uploaded to web for the pre-trial application, and it was applied with seven undergraduate students to see if it has a clear and understandable form. Minor revisions were made in line with the feedbacks.

Part II. The second part of the tool is primarily related with the assessment approaches used in the courses during the Covid-19 pandemic. Questionnaire consists of 7 items using 4-point likert type ranging from (1) *never used*, (2) *used in some courses*, (3) *used in most of the courses*, (4) *used in all courses*. The aim of this part is to observe what kind of approaches or techniques were preferred. In this part, the participants are also asked about whether they took online tests. Additionally, students who took the online tests were asked to mark which of the following security measures were taken in the exam.

- There was a time limit.
- The items were presented randomly to each student (order of items was unique for each student).
- Answer choices were presented randomly (order of choices was unique for each student).
- Different items were used (there was an item pool).
- Cameras were required to be open during the exams.
- There was control not to allow opening a new web page/tab.

Part III. In the third and last part of the data collection tool, participants' views about the assessment practices were aimed to be determined. In this section, there are open-ended questions that investigate the participants' views about assessment tools used in distance education, about uncovering the assessment preferences of the participants, and comparing face-to-face and distance assessment practices, whether participants experienced technical problems, and revealing participations' communication level with the instructors.

Volunteerism is of great importance for two main reasons; the accuracy of the data and the potential to threat validity and reliability of the instrument. Informed consent form is included on the first page of the e-form to ensure that only volunteers are included as participants. The data were collected in approximately one and a half month with efforts of the researchers using all of their social networks. Because of low number of returns to online surveys, the total number of participants could only reach 486.

2.3. Data Analysis

The procedure followed in the analysis of the data are as follows according to the parts of data collection tool and research purposes. MS Excel and IBM SPSS 20 were used for analysis.

1. Since researchers aimed to measure the participants' perceptions about the quality of assessment practices, first part of the instrument was developed as a measurement tool and explanatory factor analysis (EFA) was conducted for reliability.
2. Descriptive statistics were calculated for the data obtained from Part I, Part II and Part III. Frequencies, percentages, total and average points, and standard deviation score were calculated. The findings were plotted so that the results can be easily understood by the reader.
3. Content analysis was conducted for qualitative data. Qualitative data was collected through answers given by the students to the open-ended question located in the Part III of the data collection tool. Details about the trustworthiness of the qualitative analysis were presented separately.
4. One-way ANOVA and independent samples t-test were conducted to answer the fourth research question. Kolmogorov-Smirnov test results ($p > .05$) indicated that data is normally distributed and Levene test showed homogeneity of variances is achieved ($p > .05$).

2.3.1. Construct Validity and Reliability of Instrument

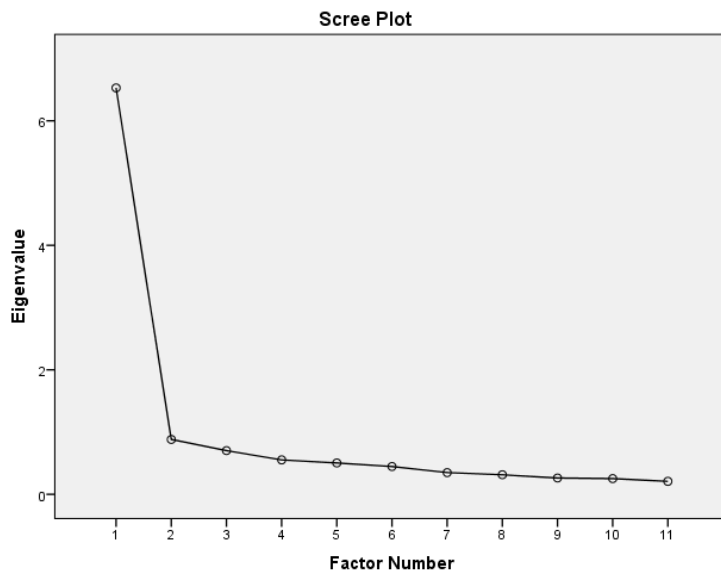
First part of the data collection tool was an *instrument* that measures *higher education students' perceptions about the quality of the assessment practices*. The instrument has 11 items and the highest score that can be obtained from the instrument is 55, and the lowest score is 11. EFA was performed using principal axis factoring method to determine the psychometric properties of instrument. First, researchers examined whether there are one-dimensional/multi-dimensional outliers in the data. It was observed that there are no extreme values. Second, sampling adequacy for EFA was examined. Kaiser-Meyer-Olkin test statistic was found as 0.932 which means perfect sample adequacy for EFA. Third, Bartlett's test of sphericity was used to investigate the multivariate normality. Results ($X^2(66) = 3454.236; p < .01$) indicated that multivariate normality was achieved. EFA results showing item factor loads are presented in [Table 3](#).

Result of the EFA presented that factor loadings of each item are between 0.648-0.842 and are gathered in one dimension. Therefore, researchers decided to use all of the items. Eigenvalue Scree Plot ([Figure 1](#)) indicates that items measure only one dimension.

Table 3. Item factor loadings.

Item Number	Items	Factor Loadings
i1	Instructions and explanations in assessment / assignments were understandable and clear.	.759
i2	I have been informed about evaluation and scoring (rubric, evaluation criteria, etc.).	.708
i3	The techniques used in assessment (homework, portfolio, open-ended questions, tests, etc.) were appropriate for the skills desired to be acquired in the lessons.	.842
i4	Assessment was aimed to measure high level skills (creative thinking, critical thinking, problem solving, etc.).	.769
i5	The effectiveness of learning was increased by rapid assessment and giving feedback.	.814
i6	Assessment results and feedback were instant.	.761
i7	The feedback was detailed and instructive.	.800
i8	Assessment practices did not allow cheating and plagiarism.	.669
i9	The assessment results were reliable.	.735
i10	Distinctiveness of test results are high.	.656
i11	The scope of the assessment did not go beyond the provided content.	.648

Figure 1. EFA Eigenvalue scree plot.



As can be seen in Figure 1, Eigenvalue of the one-dimension is calculated as 6.529. The variance explained by the one-dimension is found as 55.43%. The Cronbach Alpha internal consistency coefficient of the instrument was calculated as 0.93. The Turkish form of the instrument is provided in Appendix.

2.3.2. Trustworthiness

While validity and reliability are used for accuracy of quantitative research, trustworthiness have the same meaning for qualitative study (Guba & Lincoln, 1994). There are some strategies that must be considered like inter-coders agreement, triangulation, peer review, debriefing and rich description (Marshall & Rossman, 2014). Researchers used inter-coder agreement and rich description to provide trustworthiness of the qualitative part of this study.

Two other coders were appointed to provide inter-coder agreement. The first coder is an assistant professor and has Ph.D. degree in measurement and evaluation. Second coder is a Ph.D. student experienced in qualitative methods. Four coders met and discussed the procedure of the study prior to coding and coded six units of data. Researcher and coders compared their findings and negotiated on differences and agreed on codes. After all coding is completed, inter-coder agreement between four coders is found as .88, as Miles and Huberman (1994) reported .80 inter-coder reliability score is satisfying. Researchers and coders compared their findings, negotiated on differences and agreed on results.

Rich description is the second strategy that researcher used for the trustworthiness of qualitative part. Researchers must indicate in-depth information about the procedure and steps of the qualitative phase of the study. The aim of detailed explanation is to provide easy understanding of phases and results (Marshall & Rossman, 2014). The researcher gave details of the qualitative phase to provide rich description so that those who wish to benefit from this research may easily understand the procedure, phases, and findings.

3. RESULTS/ FINDINGS

In this section, findings related to research questions will be presented. Four sub-headings were created for four research questions.

3.1. Participants' Perceptions about the Quality of Assessment Practices

The descriptive statistics regarding the participants' perceptions about the quality of the assessment practices carried out during the Covid-19 pandemic process are presented in [Table 4](#).


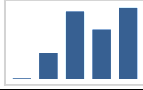

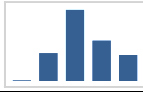
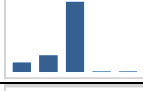






Table 4. Descriptives of participants' perceptions about the quality of assessment practices.

Range	Minimum	Maximum	Mean	S.D.	Variance	Skewness	Kurtosis
44.00	11.00	55.00	35.29	11.00	121.01	-.160	-.51

As can be noticed in [Table 4](#), skewness and kurtosis values prove participants' perceptions scores about the quality of assessment practices is normally distributed. However, it can be said that it is skewed to left although not at a significant level. This means big part of the observations are medium/large, with a few observations that are much smaller. As a matter of fact, the distance of the average ($\bar{X} = 35.29$) is closer to maximum score than the lowest score. This presents a clue about the participants' perceptions tend to be relatively moderate to high. However, since this is not a statistically significant distortion, it can be stated that participants' perceptions of the quality of assessment practices are moderate. The distribution of the responses, the average and standard deviation values for each item are presented in [Table 5](#).

According to [Table 5](#), the average of all items except two items (i5 and i10) are found as 3.00 and above. Results show that clarity of the instructions and explanations used in assessment practices are high ($\bar{X} = 3.56$, $S = 1.23$). On the other hand, participants negatively valued about the use of instant assessment and feedback. In other words, instant assessment and giving feedback ($\bar{X} = 2.94$, $S = 1.34$) are not sufficient to increase the effectiveness of the participants' learning. In addition, participants think that the test results do not have enough power to distinguish the students ($\bar{X} = 2.73$, $S = 1.35$).

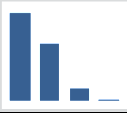
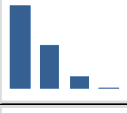

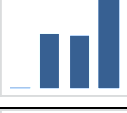



Table 5. Participants’ perceptions about the quality of assessment practices.

No	Item	definitely disagree	disagree	partly agree	agree	definitely agree	Mini Graph	\bar{X}	S
i1	Instructions and explanations in assessment / assignments were understandable and clear.	41	49	124	139	133		3.56	1.23
i2	I have been informed about evaluation and scoring (rubric, evaluation criteria, etc.).	48	78	125	105	130		3.39	1.30
i3	The techniques used in assessment (homework, portfolio, open-ended questions, tests, etc.) were appropriate for the skills desired to be acquired in the lessons.	43	75	150	108	110		3.34	1.23
i4	Assessment was aimed to measure high level skills (creative thinking, critical thinking, problem solving, etc.).	60	92	139	105	90		3.15	1.27
i5	The effectiveness of learning was increased by rapid assessment and giving feedback.	90	95	136	82	83		2.94	1.34
i6	Assessment results and feedback were instant.	56	83	133	102	112		3.27	1.30
i7	The feedback was detailed and instructive.	77	96	143	92	78		3.00	1.29
i8	Assessment practices did not allow cheating and plagiarism.	80	80	108	88	130		3.22	1.42
i9	The assessment results were reliable.	73	75	132	113	93		3.16	1.31
i10	Distinctiveness of test results are high.	122	97	122	80	65		2.73	1.35
i11	The scope of the assessment did not go beyond the provided content.	45	56	129	117	139		3.51	1.27

3.2. Assessment Approaches Used During the Pandemic

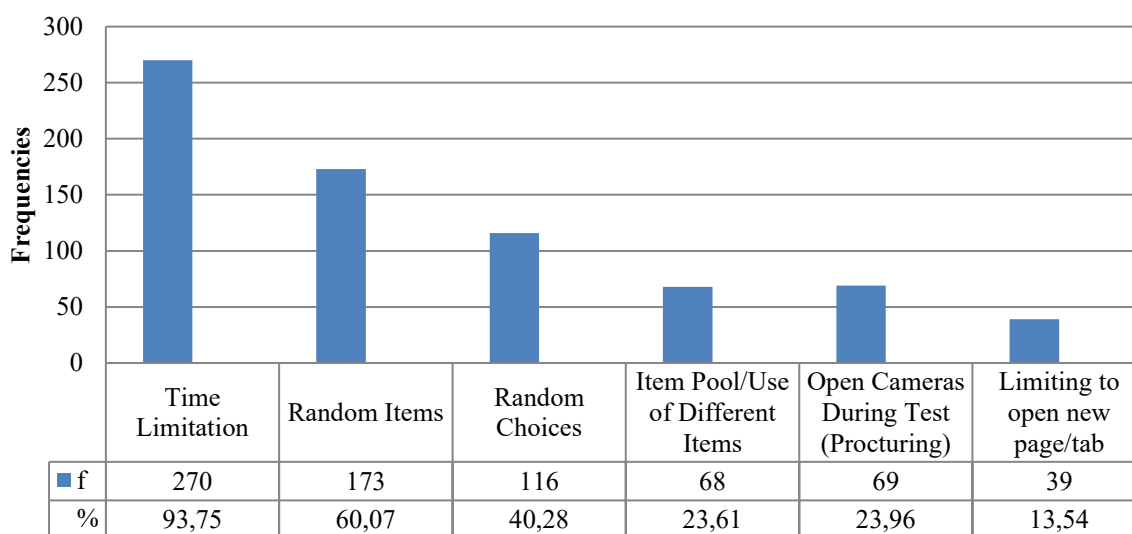
Findings regarding the usage measures of assessment approaches applied in courses during COVID-19 pandemic are presented in Table 6. As can be seen in Table 6, assignments, one of the performance-based tools, is the mostly used approach (total = 1435) overall. Assignments may be used with different techniques and in different forms. Using various approaches together such as projects, portfolios, open-ended items is the second mostly preferred approach (total = 1262). These results indicate that performance-based techniques such as open-ended questions, take home exams, product files or performance tasks were widely preferred during the pandemic. Infrequent use of online tests, peer and self assessment tools and participation indicators is an eye-catching result. As visualized in mini graphs, usage measures of these approaches are generally reported as “never used”.

Table 6. Usage measures of assessment approaches.

No	Item	never used	in some courses	in most courses	in all courses	Mini Graph	Total
1	Online tests using items that require selection (multiple choice, T-F, matching)	205	153	76	52		947
2	Online tests using open-ended (written) items	187	135	91	73		1022
3	Online tests using a combination of selection and open-ended items	243	122	68	53		903
4	Assignments with specified time (e.g., 1 week) (open-ended questions, take-home exams, or performance-based techniques such as portfolios, performance tasks)	43	128	124	191		1435
5	Various assessment techniques were used for evaluation (portfolio, research project, open-ended items etc.)	76	159	136	115		1262
6	Peer and / or self-assessment tools	154	175	91	66		1041
7	Discussion forums or other indicators showing participation in distance education	181	158	75	72		1010

Since test security is a problematic issue in remote assessment, participants who attended online tests were asked about the test security measures. 198 (40.74%) of the 486 students stated that they did not take an online test. Figure 2 summarizes the measures taken for the test security during the online tests.

Figure 2. Descriptives about the test security measures in online tests.



According to Figure 2, the most preferred test security measure in online tests is using time limitations (93.75%). The second and third mostly used measures is to present items and choices randomly (in different order) (60.07%, 40.28%) for each participant. Use of item pool/providing different questions or using open camera (procturing) are less preferred measures (23%) in online tests. Limiting to open a new web page/tab is the least preferred security measure (13.54%). Participants were asked to declare which other security measures they experienced. Replies of the participants were listed as follows.

- Recording a video narration explaining answers.
- Asking too many items in limited time (e.g. 90 items 20 minutes).
- Limiting the monitor/control of the responded items.

3.3. Views of Students on Assessment Practices

Third part of the data collection tool was aimed to identify views of participants about assessment practices. Both quantitative and qualitative data were collected. Four 5-point Likert type items were presented for the quantitative part. Table 7 summarizes the replies of the participants. According to Table 7, the highest average score ($\bar{X}_4 = 3.76$) is related with the interaction of faculty members and students. The second highest average score ($\bar{X}_1 = 3.24$) is about the test anxiety. Despite being positive about the quality of the assessment practices (first research question) and not having technical problems, participants reported that they are highly concerned about remote assessment. Similarly, participants are not likely to prefer remote assessment when face-to-face education begins. Although the low number of participants experienced technical problems is a pleasing finding indicating sufficient infrastructure of distance learning systems, the fact that even a student is experiencing a technical problem may indicate an important problem that will question the validity of the scores and prevent fair measurement.

Table 7. Participants' views on assessment practices.

No	Item	definitely disagree	disagree	partly agree	agree	definitely agree	Mini Graph	\bar{X}	S
1	I was more concerned about remote assessment than I feel in face-to-face assessment	90	75	100	71	150		3.24	1.49
2	I prefer remote assessment when face-to-face education begins.	180	59	107	50	90		2.61	1.52
3	I had technical problems in sending assignments/tests etc.	195	99	86	53	53		2.32	1.38
4	I could contact instructor when I had questions about assignments	30	58	100	111	187		3.76	1.25

In the last part of the data collection tool, participants were asked whether they would like to state their views about the assessment practices carried out during COVID-19 pandemic. 175 of the participants answered this part. Using content analysis, codes were grouped into the categories as negative views, positive views and demands of the participants. Codes and frequencies were given in Table 8. Additionally, it was observed that, apart from the focus of this study, participants are inclined to state views comparing face-to-face and distance education.

Table 8. *Frequencies of codes.*

Demands	<i>f</i>	Negative Views	<i>f</i>	Positive Views	<i>f</i>
Use of assignments	14	Distinctiveness of scores	11	Independent from time and place	6
Use of online tests	9	Items/assignments out of content	12	Aimed to measure high level skills	4
Interaction and feedback	8	Time limitations	12	Not having exam anxiety	4
Content of the tests/assignments	6	Negligence of the evaluaters	7	Interaction and feedback	2
Use of clear exam/assignment instructions	5	Use of online tests	6		
Use of varied assessment practices	5	Overrated scores	4		
Use of rubrics	4	Limited interaction and feedback	4		
Technical infrastructure	4	Limited Measurement of high-level skills	4		
Use of face-to-face exams	3	Technical problems	4		
Measuring high level skills	3	Lack of clear exam/assignment instructions	4		
Individualized assessment	3	Lack of clear instructions	2		
Test security	2				

Participants highly reported that they demand to use assignments and online tests for assessment. Another demand of the students is about the interaction and feedback. Since distance education do not offer classroom environment, student-student and student-faculty member interaction is getting more importance (Alhih & Ossiannilsson, 2017). Participants also reported negative views. The mostly declared negative view is about the distinctiveness of the scores. There is a common view among students that most there are excessively overrated scores. This may be reasoned from the heavy workload of the faculty members since all of the courses are given online and there was plenty of assignments to mark. Participants also declared negative views about “content of the tests/assignments” and “time limitations”. Participants highly criticized the exams and assignments since they think that content is extensive, faculty members demanded assignments whose subject is out of course content and there are strict time limitations, especially for assignments. Participants’ declared positive views about the time and place independence that distance education presents, aim of measuring higher level skills and exam anxiety but all of them are limited.

3.4. Quality of assessment according to Level of Interaction with Faculty Members and Taking Online Tests

Literature offers strong relationship between interaction and students’ perception in distance education. In this study, researchers decided to examine if there is any significant difference in participants’ perceptions about the quality of assessment practices according to level of interaction with faculty members. Participants were grouped according to their reply one of the items (Item 4 - *I could contant to instructor when I had questions about assingments*) in the third part of the data collection tool. Descriptives are provided in [Table 9](#).

Table 9. Participants' replies to Item 4.

No	Item4- I could contact instructor when I had questions about assignments	f	%
1	Strongly disagree	30	6.17
2	Disagree	58	11.93
3	Partly agree	100	20.58
4	Agree	111	22.84
5	Strongly agree	187	38.48

One-way ANOVA was employed, and five groups of participants were compared. Results are presented in Table 10.

Table 10. ANOVA results on perception of assessment quality * interaction level with faculty members.

	Sum of Squares	df	Mean Square	F	Sig.	Sig. Dif.
Between Groups	22149.249	4	5537.312	72.891	.000	1-2; 1-3;1-4;1-5;
Within Groups	36539.995	481	75.967			2-4; 2-5; 3-4; 3-5;
Total	58689.245	485				4-5

ANOVA results proved that participants' perceptions about the quality of assessment practices differ significantly according to participants' level of interaction with faculty members, $F(4, 481)=72.861, p<.01$. There is significant difference ($p<.01$) between all levels of participants except *Disagree* and *Partially Agree* groups.

Online tests which are widely used in remote assessment has problems in test security. On the other hand, the qualitative phase of this study reported that students support the use of online tests. With these in mind, we decided to examine whether the use of online tests effect participants' perception about quality of assessment practices. Table 11 summarizes the results of independent samples *t*-test.

Table 11. Results of independent samples *t*-test.

Groups	n	\bar{X}	S	df	t	p
Attended Online Tests	288	40.48	11.88	484	3.26	.001
Not Attended Online Tests	198	36.48	11.56			

Results of independent samples *t*-test indicated participants' perceptions about the the quality of the assessment practices are significantly different according to participants' attendance to online tests, $t(484)=3.26; p<.01$. Participants who took online tests ($\bar{X}=40.48$) have higher perceptions about the quality of the assessment practices than participants who did not attend online tests ($\bar{X}=36.48$).

4. DISCUSSION and CONCLUSION

As badly affected all the routines, pandemic changed the way we teach. While we had theoretical definitions and limited practices of distance education earlier, nowadays, distance education has a meaning for all. Today we use online tools to make remote lessons, to communicate and interact, to assign and collect homeworks and conduct assessment. In this research, we aimed to examine the very first use of remote assessment and participants' views about this unique experience asking four research questions.

First of all, participants reported positively about the quality of remote assessment. However, they reported negatively in two critical items. First, participants agreed that the use of rapid assessment and feedback was insufficient for effective learning. Instant feedback is known as an assistant to distance learners to self-evaluate their learning and increase performance in summative assessment (Koneru, 2017). Similarly, rapid assessment, is critical in distance education courses due to the asynchronous nature of these courses and additional effort was required to confirm that students were ready to receive and respond to feedback properly (Uribe & Vaughan, 2017). The second issue that participants negatively reported about the quality of the assessment practices is distinctiveness of test scores. To explain, students believe that assessment must produce fair test results. Most of the faculty members have experienced remote assessment tools for the first time and this may be a reason for students to feel that distance assessment practices did not yield distinctive test scores. Heavy workload can be pointed as another reason for unfair results. Faculty members gave all of the courses online and they must evaluate plenty of assignments, projects, and other remote assessment tools.

Results indicated that performance-based tools like assignments, performance tasks, portfolios and research projects are the mostly used assessment tools. Online tests which are easy-to-use were found to be used less. Participation to discussion forums or other indicators of participation rates to distance education are used infrequently, too. Additionally, participants reported infrequent use of peer or self assessment tools. However, literature offers to use varied tools for remote assessment (Stödberg, 2012). Limited use of online tests may have resulted from the concern about failing to meet test security. On the other hand, infrequent use of discussion forums or other indicators of participation to distance tools may be because of the inexperience of instructors about remote tools since most of the faculty members used these tools for the first time.

Even though participants did not have any technical problems and they have easy access to faculty members, they supported to use conventional exams rather than remote assessment. Moreover, participants reported that they did not experience anxiety during remote assessment. This may be explained by one of the findings of qualitative phase of the study. Students reported negative views about distinctiveness of the results, and this may be routing participants in favour of face-to-face assessment.

Further information about participants' views on remote assessment was aimed with qualitative data. Views of participants were grouped as positive views, negative views and demands. Participants declared negative views about the distinctiveness of assessment results. We know that assignments are the mostly used tool during pandemic according to the results of first research question and use of performance based tools like assignments, portfolios or projects may be laborious for faculty members (Linn et al., 1991) and this may lead to unfair assessment results. Another negative view is about the items/assignments that are out of content. With the use of online tools, a wider course content may be presented to students with the idea of having more self-studying time in distance education. Lastly, students may need more time for fulfilling performance-based tasks which requires process-oriented workload. Participants have demands, too. First of all, they demand the use of assignments and online tests and needs more interaction and feedback. Additionally, they demand well-defined exam/assignment instructions. Time and place independence, measuring higher level skills and lower level of exam anxiety are found as the positive sides of remote assessment.

Another finding is that participants who have higher levels of interaction with instructors find assessment practices more qualified. In other words, the more students can reach the faculty members and communicate, the more qualified they find the assessment. Student-student and students-instructor interaction or communication is critical in distance education since there are no conventional classrooms (Vlachopoulos & Makri, 2019). A similar result is found when

online tests and participants' perceptions about quality about assessment tools are examined. Participants who take online tests valued assessment tools more qualified. This may stem from that online test takers may feel a real assessment experience through online tests. As reported earlier, most of the students experienced distance education for the first time and they may need to involve a similar assessment tool as in conventional classrooms.

As all studies have, this study also has limitations, too. Although it is aimed a larger study group, only 486 students volunteered to participate in the study. Participants are from 61 different universities and 69 different faculties, but a larger group may yield detailed results. To overcome this limitation, a similar research with a larger group may be conducted. This study focused on students. Faculty members are the practitioners and their views about this phenomenon may help us to develop remote assessment approaches. This study was conducted considering the early stages of the pandemic. Covid-19 is still threatening the face-to-face education and institutions are now experienced in distant education. Future studies may focus on the developments and the latest techniques that institutions used for assessment.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the authors. **Ethics Committee Number:** Balikesir University Ethics Committee, 2020-8.

Authorship contribution statement

Selma Senel: Investigation, Development of Data Collection Tool, Visualization, and Analysis.

Huseyin Can Senel: Methodology, Analysis, Supervision, and Writing the original draft.

ORCID

Selma ŞENEL  <https://orcid.org/0000-0002-5803-0793>

Hüseyin Can ŞENEL  <https://orcid.org/0000-0002-7501-9174>

5. REFERENCES

- Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers and Education*, 126, 334-345. <https://doi.org/10.1016/j.compedu.2018.07.021>
- Alhih, M., & Ossiannilsson, E. (2017). Levels of interaction provided by online distance education models. *EURASIA Journal of Mathematics Science and Technology Education*, 13(6), 2733-2748. <https://doi.org/10.12973/eurasia.2017.01250a>
- Araka, E., Maina, E., Gitonga, R., & Oboko, R. (2020). Research trends in measurement and intervention tools for self-regulated learning for e-learning environments-systematic review (2008-2018). *Research and Practice in Technology Enhanced Learning*, 15(6), 1-21. <https://doi.org/10.1186/s41039-020-00129-5>
- Bozkurt, A. & Sharma R. C. (2020). Emergency remote teaching in a time of global crisis due to CoronaVirus pandemic. *Asian Journal of Distance Education*, 15(1), 1-6.
- Arnold, I. J. M. (2016). Cheating at online formative tests: Does it pay off? *Internet and Higher Education*, 29, 98-106. <https://doi.org/10.1016/j.iheduc.2016.02.001>
- Atılğan, H., Kan, A., & Doğan, N. (2009). *Eğitimde ölçme ve değerlendirme [Measurement and evaluation in education]*. Anı Yayıncılık.
- Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability*, 21(1), 33-46. <https://doi.org/10.1007/s11092-008-9064-9>
- Chaiyo, Y., & Nokham, R. (2017). The effect of Kahoot, Quizizz and Google Forms on the student's perception in the classrooms response system. *2nd Joint International*

- Conference on Digital Arts, Media and Technology 2017: Digital Economy for Sustainable Growth, ICDAMT 2017*, 178-182. <https://doi.org/10.1109/ICDAMT.2017.7904957>
- Chen, Z., Jiao, J., & Hu, K. (2020). Formative assessment as an online instruction intervention. *International Journal of Distance Education Technologies*, 19(1), 1-16. <https://doi.org/10.4018/ijdet.20210101.oa1>
- Cheng, I., & Basu, A. (2006). Improving multimedia innovative item types for computer based testing. *ISM 2006 - 8th IEEE International Symposium on Multimedia*, 557-564. <https://doi.org/10.1109/ISM.2006.92>
- Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects*, 49(1), 91-96. <https://doi.org/10.1007/s11125-020-09464-3>
- Frey, B. (2018). *The SAGE encyclopedia of educational research, measurement, and evaluation*. SAGE Publications. <https://doi.org/10.4135/9781506326139>
- Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). SAGE Publications.
- Hatzipanagos, S., & Warburton, S. (2009). Feedback as dialogue: Exploring the links between formative assessment and social software in distance learning. *Learning, Media and Technology*, 34(1), 45-59. <https://doi.org/10.1080/17439880902759919>
- Jain, A. K., Bolle, R., & Pankanti, S. (2006). *Biometrics: personal identification in networked society* (Vol. 479). Springer.
- Joint Information Systems Committee. (2010). *Effective assessment in a digital age a guide to technology-enhanced assessment and feedback*. Higher Education Funding Council for England. https://facultyinnovate.utexas.edu/sites/default/files/digiassass_eada.pdf
- Koh, J. H. L., & Kan, R. Y. P. (2020). Perceptions of learning management system quality, satisfaction, and usage: Differences among students of the arts. *Australasian Journal of Educational Technology*, 36(3), 26-40. <https://doi.org/10.14742/AJET.5187>
- Koneru, I. (2017). Exploring moodle functionality for managing Open Distance Learning e-assessments. *Turkish Online Journal of Distance Education*, 18(4), 129-141. <https://doi.org/10.17718/tojde.340402>
- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20(8), 15-21.
- Marshall, C., & Rossman, G. B. (2014). *Designing qualitative research*. SAGE Publications.
- Martini, M., Gazzaniga, V., Bragazzi, N. L., & Barberis, I. (2019). The Spanish influenza pandemic: A lesson from history 100 years after 1918. *Journal of Preventive Medicine and Hygiene*, 60(1), E64-E67. <https://doi.org/10.15167/2421-4248/jpmh2019.60.1.1205>
- Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). E-learning, online learning, and distance learning environments: Are they the same? *Internet and Higher Education*, 14(2), 129-135. <https://doi.org/10.1016/j.iheduc.2010.10.001>
- Murray, M., Pérez, J., Geist, D., & Hedrick, A. (2012). Student interaction with online course content: Build it and they might come. *Journal of Information Technology Education: Research*, 11(1), 125-140. <https://doi.org/10.28945/1592>
- Nguyen, J. G., Keuseman, K. J., & Humston, J. J. (2020). Minimize online cheating for online assessments during COVID-19 Pandemic. *Journal of Chemical Education*. <https://doi.org/10.1021/acs.jchemed.0c00790>
- Nyachwaya, J. M. (2020). Teaching general chemistry (I) online during COVID-19. Process, outcomes, and lessons learned: A reflection. *Journal of Chemical Education*, 1, 17-21. <https://doi.org/10.1021/acs.jchemed.0c00891>
- Peterson, J. (2019). An analysis of academic dishonesty in online classes. *Mid-Western Educational Researcher*, 31(1), 24-36.

- Rovai, A. P. (2000). Online and traditional assessments: What is the difference? *Internet and Higher Education*, 3(3), 141-151. [https://doi.org/10.1016/S1096-7516\(01\)00028-8](https://doi.org/10.1016/S1096-7516(01)00028-8)
- Rowe, N. C. (2004). Cheating in online student assessment: Beyond plagiarism. *Online Journal of Distance Learning Administration*, 7(2). 1-8.
- Shrago, J. B., & Smith, M. K. (2006). Online assessment in the K-12 classroom: A formative assessment model for improving student performance on standardized tests. In M. Hricko & L. S. Howell (Eds.), *Online assessment and measurement: Case studies from higher education, K-12 and corporate* (pp. 181-195). Information Science Publishing. <https://doi.org/10.4018/978-1-59140-497-2.ch013>
- Stödberg, U. (2012). A research review of e-assessment. *Assessment and Evaluation in Higher Education*, 37(5), 591-604. <https://doi.org/10.1080/02602938.2011.557496>
- University Ranking by Academic Performance. (2020, November). *The position of our universities in the 11 world general rankings in 2019*. <http://tr.urapcenter.org/2019/index.php>
- Uribe, S. N., & Vaughan, M. (2017). Facilitating student learning in distance education: a case study on the development and implementation of a multifaceted feedback system. *Distance Education*, 38(3), 288-301. <https://doi.org/10.1080/01587919.2017.1369005>
- Veletsianos, G., & Houlden, S. (2019). An analysis of flexible learning and flexibility over the last 40 years of distance education. *Distance Education*, 40(4), 454-468. <https://doi.org/10.1080/01587919.2019.1681893>
- Vlachopoulos, D., & Makri, A. (2019). Online communication and interaction in distance higher education: A framework study of good practice. *International Review of Education*, 65(4), 605-632. <https://doi.org/10.1007/s11159-019-09792-3>
- Williams, D. D., Howell, S. L., & Hricko, M. (2005). *Online assessment, measurement and evaluation: Emerging practices*. Information Science Publishing. <https://doi.org/10.4018/978-1-59140-747-8>

6. APPENDIX

Instrument for Student Perceptions About the Quality of the Assessment (Turkish Form)

[Ölçme ve Değerlendirmenin Niteliğine İlişkin Öğrenci Algısı Ölçeği]

Bu ölçekte, uzaktan eğitim sürecinde karşılaşmış olduğun ölçme ve değerlendirme işlemlerine ilişkin algının belirlenmesi amaçlanmaktadır. Maddelerin her birini okuyarak, “*Kesinlikle Katılmıyorum, Katılmıyorum, Kısmen Katılıyorum, Katılıyorum, Kesinlikle Katılıyorum*” seçeneklerinden birini işaretlemeniz beklenmektedir. Araştırmaya desteğinizden dolayı teşekkür ederiz.

No	Maddeler	Kesinlikle Katılmıyorum	Katılmıyorum	Kısmen Katılıyorum	Katılıyorum	Kesinlikle Katılıyorum
i1	Ölçme/ödevlendirme sürecindeki yönerge ve açıklamalar anlaşılır ve açıktı.					
i2	Değerlendirme ve puanlamanın nasıl yapılacağı konusunda bilgilendirildim (dereceli puanlama anahtarı, değerlendirme kriterleri vb.).					
i3	Ölçmede kullanılan teknikler (ödev, ürün dosyası, açık uçlu soru, test vb.) derslerde kazandırılmak istenen becerilere uygundu.					
i4	Ölçme ve değerlendirme üst düzey becerileri (yaratıcı düşünme, eleştirel düşünme, problem çözme vb.) yoklar nitelikteydi.					
i5	Eğitim süreci boyunca ölçme yapılarak, dönütler verilerek öğrenme sürecimin etkililiği arttırıldı.					
i6	Ölçme sonuçları ve dönütler hızlıca ulaştı.					
i7	Geribildirimler ayrıntılı ve öğreticiydi.					
i8	Ölçme ve değerlendirme kopya ve intihale (farklı kaynaklardan kaynak göstermeden alma) izin vermeyecek biçimde yapıldı.					
i9	Ölçme sonuçları güvenilirirdi (hatasızdı).					
i10	Sınav sonuçlarının başarılı ve başarısız ayırt ediciliği yüksekti.					
i11	Ölçme kapsamı, sunulan ders içeriği dışına çıkmadı.					