

SELF-ASSESSMENT QUIZ TAKING BEHAVIOUR ANALYSIS IN AN ONLINE COURSE

Yasin Ozarslan [yasin.ozarslan@yasar.edu.tr], Ozlem Ozan [ozlem.ozan@yasar.edu.tr],
Yasar University [http://yasar.edu.tr], Turkey

Abstract

Self-assessment is vital for online learning since it is one of the most essential skills of distance learners. In this respect, the purpose of this study was to understand learners' self-assessment quiz taking behaviours in an undergraduate level online course. We tried to figure out whether there is a relation between self-assessment quiz taking behaviours and final exam scores or not. In addition, we investigated how self-assessment quiz taking behaviour differs with respect to learner profile. In line with this purpose, 677 students' 6092 test events across Project Culture course on Sakai CLE LMS were analyzed. For the analysis of the quantitative data, one-way ANOVA, Chi-Square test of independence, independent-samples t-test and descriptive statistics were utilized. The results revealed that learners who attended self-assessment quizzes regularly had higher final exam scores than others who did not attend those quizzes. Also, they were more satisfied with the course than others study field. In addition, learners who attended self-assessment quizzes regularly had a higher degree of perceived learning. However, number of attempts to those quizzes does not have an effect on final exam scores. On the other hand, a statistically significant relationship was found between attempt number and gender in favour of female learners.

Keywords: Self-Assessment, Self-Assessment quiz, test event analysis, online course, final exam scores, course grade

Abstract in Turkish

Öz-değerlendirme, uzaktan öğrenen için en temel becerilerden biri olduğundan çevrimiçi öğrenmede için çok önemlidir. Bu bağlamda, bu çalışmada öğrenenlerin kısa öz-değerlendirme testlerine katılma davranışları araştırılmıştır. Öz-değerlendirme testlerine katılma durumunun ders başarısında bir farklılaşmaya sebep olup olmadığı ve öğrenen profiline göre nasıl değiştiği incelenmiştir. Çalışma bağlamında Sakai CLE LMS üzerinden sunulan Proje Kültürü dersi kapsamında 677 öğrencinin 6092 test olayı analiz edilmiştir. Veri analizinde Tek yönlü varyans analizi, Chi-Square testi, bağımsız örneklem t-testi ve betimsel istatistikler kullanılmıştır. Araştırma sonuçlarına göre öz-değerlendirme testlerine katılan öğrencilerin ders başarıları katılmayanlara göre daha yüksektir. Ayrıca, öz-değerlendirme testlerine düzenli olarak katılan öğrencilerin ders memnuniyeti ve algılanan öğrenme düzeyleri katılmayanlara göre daha yüksektir. Öz-değerlendirme testlerini tekrarlama sayısı ile ders başarısı arasında anlamlı bir ilişki gözlemlenmemiştir. Öz-değerlendirme testlerine katılma davranışı, öğrenenin cinsiyeti ve bölümüne göre farklılık göstermemektedir. Öte yandan öz-değerlendirme testlerini tekrar yapma yüzdesi kadın öğrenenler lehine anlamlı bir farklılık göstermektedir.

Introduction

Self-assessment is useful in providing learners with better understanding of the subject matter (Claxton, 1995). It leads to motivation (Pintrich & Schunk, 2002; Schunk, 2004), enhance learning (Boud, 2000), improvement of results in the final exam (Cassady & Gridley, 2005; Ćukušić, et al., 2014; Ibabe & Jauregizar, 2009; Wilson, Boyd, Chen, & Jamal, 2011), beneficial for improvement of learner engagement (Gikandi, et al., 2011).

The question addressed in this study is whether the self-assessment quizzes would form any differentiation in learning that we can detect from log files. We tried to find out what patterns in self-assessment quiz taking could be detected in log files. We focused on the information (test event data, learner profile), which we could acquire from log files and a mini survey, which is about the perceived learning and satisfaction with the course. We parsed the raw of log files as Number of Attempts (No Attempt, Attempt Once, and Multiple Attempts), Self-Assessment Quiz Taking Behaviour (No attendance, Irregular attendance, Regular attendance with one attempt, Regular attendance with multiple attempts) and learner profile (gender and department). After that, we asked the research questions below to detect the behaviours in self-assessment quiz taking and to see if final exam scores differ significantly with respect to those behaviours:

- Do final exam scores differ significantly with respect to self-assessment quiz taking behaviour of learners?
- Do learner satisfaction and perceived learning differ significantly with respect to self-assessment quiz taking behaviour?
- Does self-assessment quiz taking behaviour differ significantly with respect to learner profile?
- Do self-assessment test events differ significantly with respect to learner profile?

The aim of this study is to understand learners' self-assessment quiz taking behaviors in the Project Culture online course, which was delivered by one of the biggest and oldest foundation university in western Turkey in the fall of 2015, as a part of course evaluation activities to improve course quality.

Background and related work

Literature on the concept of self-assessment is quite diverse. It is not new and discussed in many instructional design textbooks (Gagne, et al., 1988; Gale, 1984; Laurillard, 1993; Rowntree, 1991; Taylor, 1998). Self-assessment is described by Oscarson, (1978) as an individual's ability to identify and self-evaluate his/her own skills in a particular area of expertise. According to Panadero (2011), "self-assessment is the qualitative assessment of the learning process, and of its final product, realized on the basis of pre-established criteria". According to McMillan and Hearn (2008) Self-assessment is a cyclic process as represented in Figure 1, and it occurs when students judge their own work to improve performance as they identify discrepancies between current and desired performance.



Figure 1. Student self-assessment cycle, (McMillan & Hearn, 2008)

As well as McMillan and Hearn (2008), Rolheiser and Ross (2013) and Zimmerman (2002) define self-assessment as learners' judgment of their own work, based on evidence and explicit criteria, for the purpose of improving future performance.

There are several characteristics of self-assessment (Heidi & Du, 2007). First, it should base on transparent criteria, which can be generated by either instructor or student or co-generated by both the instructor and the students (Dochy & McDowell, 1997; Garcia & Floyd, 1999; Frederiksen & Collins, 1989; Hanrahan & Isaacs, 2001; Wiggins, 1998; Stiggins, 2001). Second, it should provide feedback that guides students' efforts and strategies (Adams, 1998; Paris & Paris, 2001; Horner & Shwery, 2002; Taras, 2005). Third, it should be ongoing and allow learners monitoring themselves regularly (Andrade & Boulay, 2003; Goodrich, 1996; Gikandi et al, 2011). In our case, as a limitation, criteria for self-assessment were instructor-generated since allowing students to define their assessments criteria was not manageable due to the high enrolment rates (approximately 600 students in each semester). Feedback provided for each question in the self-assessment quizzes as brief information sharing. In addition, self-assessment quizzes placed in each module to provide students with ongoing monitoring of themselves.

According to Boud and Falchikov (1989), self-assessment is a "student activity through which he/she judges their own learning". Bourke (2010) states that if we look self-assessment from the viewpoint of students, the aim of it can be categorized under six types as (a) seeking an opinion of others (especially teacher) if they learnt, (b) getting marks and grades as an external measure, (c) reflecting on their performance, (d) identifying their role in learning and assessment process, (e) setting learning goals and (f) evaluating learning content. In our case, self-assessment quizzes were limited with automated versions of type 1 and 2 level due to again large number of students. Interaction with each of 677 students to provide an instantaneous reflective feedback for their self-assessment was not manageable because of the work load it could create.

Ibabe and Jauregizar (2009) distinguishes *self-testing* and self-assessment. Self-testing involves students' checking their performance against provided test items (with right and wrong answers) (Boud and Brew (1995) cited in Ibabe and Jauregizar (2009)). The self-assessment quizzes in our case could be understood as self-testing tasks, but with feedback. These kind of practice tests (Wilson et al., 2011) can improve performance (Bälter et al, 2013; Gretes & Green, 2000; Snooks, 2004), and provide students with the opportunity to review course material, demonstrate knowledge and identify weaknesses (Kulik et al., 1984), see instructor's expectations (Snooks, 2004), increase students' attitudes positively (Deutsch et al., 2012). There are also disadvantages with assessment in online environment (Bälter et al., 2013; Zhang et al., 2004). One of them is the problem of knowing who is answering the questions and also whether this person is receiving help or not (Carter et al., 2003). Another difficulty is that almost correct solutions cannot be handled as smoothly as on face-to-face environment (Bälter et al., 2013). Although many researches highlight the effectiveness of self-assessment in educational settings, as presented

above, one of them focuses on test analytics of learners' in online courses as big data (Ćukušić et al., 2014). The analysis which is shared in this paper was conducted to investigate how self-assessment quiz taking behavior occurs in our system.

Methodology

Data was gathered from user sessions of online Project Culture course, which run on Sakai CLE LMS. Test Event sessions of learners were extracted from overall course data and analyzed. Test event session represents a single instance of a student attempting a particular self-assessment test. Each test event session contains username, course code, test ID, numbers of attempts, and score. 6092 test events of 677 students were analyzed. In addition, data of perceived learning and course satisfaction gathered from a mini survey which was utilized at the end of the course.

Course selection

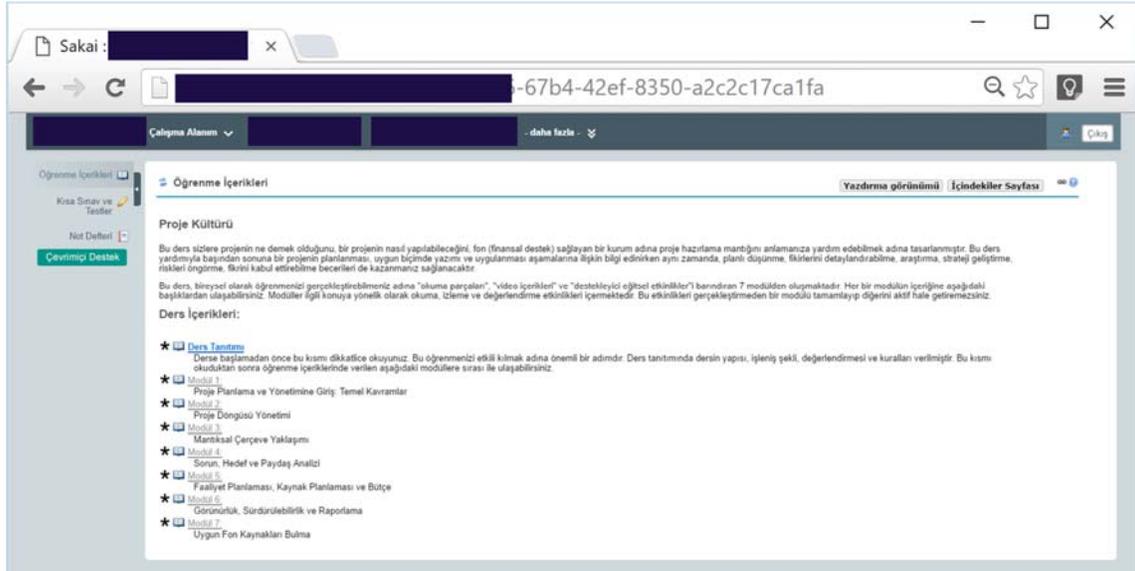
Project Culture Course was used as the case in this study. It is a foundation course of the university. This course is chosen by the students studying on Bachelor programs (Faculty of Science and Letters, Faculty of Fine Arts, Faculty of Law, Faculty of Economics and Administrative Sciences, Faculty of Communication, Faculty of Architecture and Faculty of Engineering) or on Associate Degree programs (Vocational School). It is a compulsory course for students of all departments within each faculty and vocational school program.

The aim of the course was to teach what were considered to be the basic knowledge of project design. The main point is to provide the basic information for students to perform practicable projects in the direction of their real experiences and support them with producing new projects. We preferred this course because it is an independent study field, it is neither science-centric such as calculus or physics, nor social-science-centric such as law or communication.

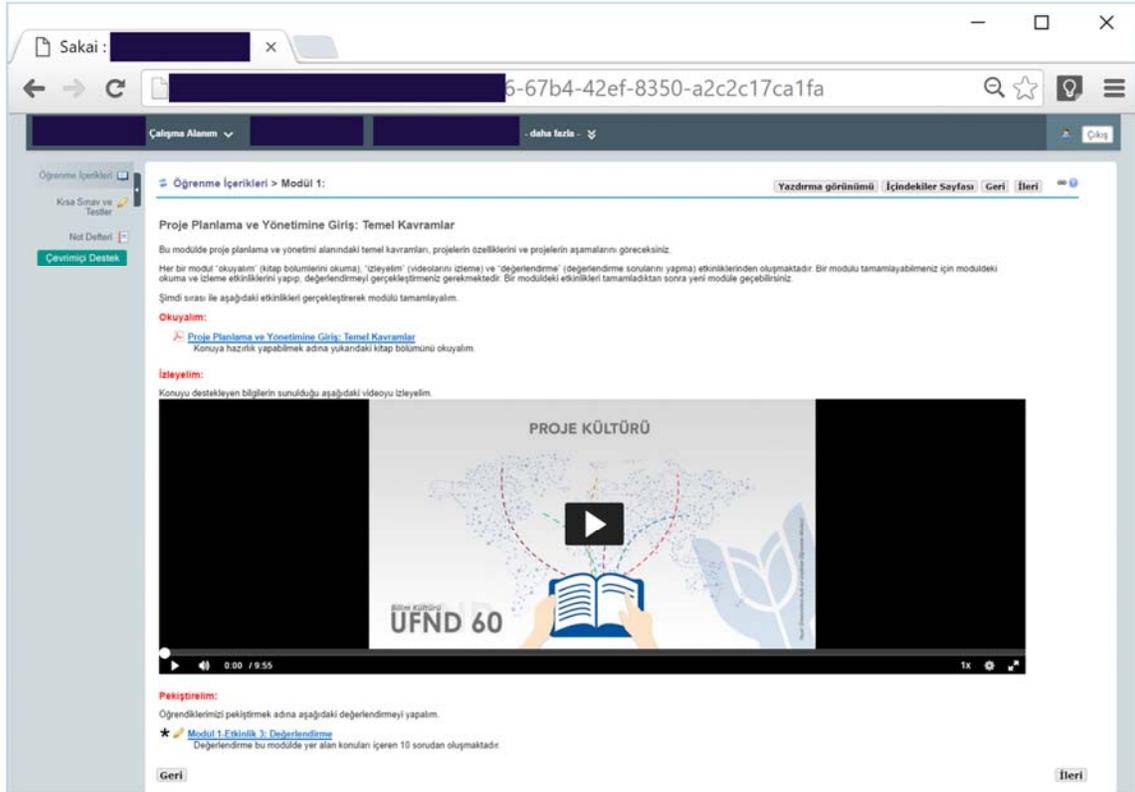
Structure and production of the course

At the beginning of the production process of course; subject matter experts, instructional designers, developers and administrators determined the content and structure of the course by getting together. It was decided for the course to include 7 modules and each module included a 50-pages reading material, 20-minutes watching material and a self assessment activity of 10 questions. In this regard, 50-pages reading material, video text and 100 questions were prepared by specialists for each module. Each module was worked on by different specialist and necessary royalty was paid to specialists. 350 pages reading material, 140 minutes video and 700 questions were produced in total for the course. After contents are prepared, they were reviewed by instructional designers and grammatical editing was provided by editors. Then, instructional design was performed and production process started. The process of design and production of the course took six months in total.

Project Culture course is composed of 7 modules. Structure of course and modules in LMS was provided in *Screenshot 1* (Course Structure), *Screenshot 2* (Course Module Structure) and *Screenshot 3* (Self-Assessment Test Page).



Screenshot 1. Course Structure



Screenshot 2. Course Module Structure

Proje Kültürü Özdeğerlendirme Modül 1

Geribildirimi Göster

İçindekiler

Bölüm 1 / 1 -

Soru 1 / 10

10.0 Puan

Projenin planlanan faaliyetlerin takvim yerine görsel bir şekilde temsili ne ile sağlanır?

- A. Gantt şeması
- B. Kritik yol ağı
- C. Ürün akış diyagramı
- D. Pareto grafiği

[Seçimi Sıfırla](#)

Ger

İleri

Kaydet

Çıkış

Screenshot 3. Self-Assessment Test Page

As mentioned earlier there was one self-assessment quiz at the end of each module. It consists of ten multiple-choice questions concerning issues raised by the related module. The quiz questions cover all of the main points of modules. For each module a question pool, that contains at least 100 questions, was created. Learner comes across 10 different random questions at each attempt. After each attempt, learner is given automated feedback. Feedback includes a short explanation about right and wrong answers and refers to related topic in module. Number of attempts is not limited. Online multiple-choice tests were used for self-assessment because of its ease of use for both developers and learners. One example for questions and feedback was provided in *Screenshot 4 and 5*. Translation of the question, is provided in Table 1:

Table 1: Translation of the sample question and feedbacks into English

Question:	What is a visual representation of a project's planned activities against a calendar called? a. A Gantt chart b. A critical path network c. A product flow diagram d. A Pareto chart
Feedback for correct answer	Congratulations, You answer the question of <i>Project Management Tools</i> chapter right.
Feedback for wrong answer	Sorry, that is incorrect! Please review <i>Project Management Tools</i> chapter in your book again.

Kısa Sınav ve Testler

60

Proje Kültürü Özdeğerlendirme Modül 1

Geribildirim Göster İçindekiler

Bölüm 1 / 1 -

Soru 1 / 10

10.0 Puan

Projenin planlanan faaliyetlerin takvim yerine görsel bir şekilde temsili ne ile sağlanır?

- ✓ A. Gantt şeması
 B. Kritik yol ağı
 C. Ürün akış diyagramı
 D. Pareto grafiği

[Seçimi Sıfırla](#)

Cevap Anahtarı: A

Geribildirim: Tebrikler, "Proje Yönetim Araçları" konu başlığına ait soruyu doğru cevapladınız.

[Geri](#) [İleri](#) [Kaydet](#) [Çıkış](#)

Screenshot 4. Correct answer

Kısa Sınav ve Testler

60

Proje Kültürü Özdeğerlendirme Modül 1

Geribildirim Göster İçindekiler

Bölüm 1 / 1 -

Soru 1 / 10

10.0 Puan

Projenin planlanan faaliyetlerin takvim yerine görsel bir şekilde temsili ne ile sağlanır?

- A. Gantt şeması
 B. Kritik yol ağı
 C. Ürün akış diyagramı
 D. Pareto grafiği

[Seçimi Sıfırla](#)

Geribildirim: Yanlış cevapladınız! Kitabınızdaki "Proje Yönetim Araçları" konu başlığını tekrar gözden geçiriniz.

[Geri](#) [İleri](#) [Kaydet](#) [Çıkış](#)

Screenshot 5. Wrong answer

Self-assessment quizzes were designed in order to provide learners with an opportunity to ascertain their mastery in particular topics and to assess their understanding of the module. The aim of those was to provide learners with the opportunity to have opinion about their learning. Attending this self-assessment activity was voluntary.

Test event data

Test event data includes username, course code, test id, test score and number of attempts.

Test score

This indicates the number of points a student earned on a self-assessment test.

Number of attempts

A test which is embedded in the course site triggers various events. The Event Logs for all tests are created and maintained automatically in Sakai LMS. Number of attempts, which is the number of retaking a self-assessment test, gathered from those records. In this study, Number of Attempts was categorized as No Attempt, Attempt Once, and Multiple Attempts. *No attempt* case occurs when a learner opens the test page but does not take the test. *Attempt once* case occurs when a learner takes the test once. *Multiple attempts* case occurs when a learner takes the test more than once.

Learner data

Learner data includes profile data and self-assessment quiz taking behaviour of learners.

Profile data of learners

This is the data about characteristics of learners who enrolled the online courses. Profile data covers final exam scores, departments, grades and genders of learners. Descriptive statistics of learner profile data was represented in Table 2.

Table 2: Number of students and achievement of students in project culture course

Gender	N (677)	%	Total Course Grade Average
Male	353	52.1	72.93
Female	324	47.9	73.24
		100	73.08

Table 3: Enrolment by faculty and vocational school program

Faculty	N	%
Faculty of Science and Letters	65	9.6
Faculty of Law	130	19.2
Faculty of Economics and Administrative Sciences	134	19.8
Faculty of Communication	58	8.6
Faculty of Architecture	79	11.7
Faculty of Engineering	122	18.0
Faculty of Art and Design	47	6.9
Vocational School	42	6.2
Total	677	100

Self-Assessment Quiz Taking Behaviour

Although taking self-assessment tests was voluntary and there was no restriction on number of attempts, some of the learners did not take any of seven self-assessment tests. Some of them took self-assessment tests of some modules. Some of them took the self-assessment tests of each module with just one attempt, and some of them took the self-assessment tests of each module

with multiple attempts. Therefore, *self-assessment quiz taking behaviour* was categorized as four types as follows:

- No attendance: This case means never taking a self-assessment test. It refers to the situation that participants did not take any of the self-assessments tests.
- Irregular attendance: This case occurs when a learner takes self-assessment test of some modules, not seven of them. It refers to the situation that participants took some of the self-assessments tests.
- Regular attendance with one attempt: This case refers to the situation that participants took all of the self-assessments tests with just one attempt in each test.
- Regular attendance with multiple attempts: This case refers to the situation that participants took all of the self-assessments tests with multiple attempts in each test.

Descriptive statistics of self-assessment quiz taking behaviour are represented in Table 4.

Table 4: Descriptive statistics of self-assessment quiz taking behaviour

Code	Self-assessment Quiz Taking Behaviour	# of Students
Type 0	No attendance	64
Type 1	Irregular attendance	268
Type 2	Regular attendance with one attempt	153
Type 3	Regular attendance with multiple attempts	192
Total		677

Survey data

A course evaluation survey, which was composed of two Likert scale questions and one open-ended question as presented in Table 5, was conducted at end of the course. The aim of course evaluation survey was to gather data about satisfaction and perceived learning of learners.

Table 5: Course evaluation survey

No	Category	Question	Strongly Disagree (1)	Disagree (2)	Neither Agree nor Disagree (3)	Agree (4)	Strongly Agree (5)
Q1	Perceived	Overall, I found this course helpful to learn the topic	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Q2	Satisfaction	I am satisfied with the course	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Q3	Comments	Do you have any suggestions or comments to improve course content?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Data analysis

For the analysis of the quantitative data, one-way ANOVA, independent samples t-test, Chi-Square test of independence and descriptive statistics were applied. For the normal distribution of data, the skewness and kurtosis coefficients were examined. First, for the normality test skewness coefficient of a distribution taken in the range of -1.5 to +1.5 and the kurtosis coefficient of a distribution taken in the range of -1.5 to +1.5 according to Tabachnick and Fidell (2013). Then, a one-way between subjects ANOVA was conducted to compare the effect of self-

assessment quiz taking behaviour on achievement. After that, an independent-samples t-test was conducted to compare self-assessment test scores and test attempt. For data of learner profile, Chi-Square test of independence was performed. Besides ANOVA, Chi-Square test of independence and independent samples t-test, descriptive statistics such as percentages (%), mean scores (\bar{X}) and frequencies (f) were used as well.

Limitations

Participants of this study were college students. They perhaps processed information differently than those without college experience would. In addition, there were no foreign students, who enrolled to the course, since the course was in Turkish. Therefore, this study should also be repeated with different participant groups and cultures.

Findings

In this section, the findings obtained in the study are presented under three headings based on research questions.

Do final exam scores differ significantly with respect to self-assessment quiz taking behaviour of learners?

A one-way between subjects ANOVA was conducted to compare the effect of self-assessment quiz taking behaviour on achievement in conditions No attendance, Irregular attendance, Regular attendance with one attempt, Regular attendance with multiple attempts. Table 6 presents the one-way ANOVA results, which demonstrate the significant effect of self-assessment quiz taking behaviour on achievement at the $p < .05$ level for the four conditions [$F(3, 673) = 4.59$, $p = 0.003$].

Table 6: One-Way Analysis of Variance of achievement

Source	df	SS	MS	F	p
Between Groups	3	1765.99	588.66	4.59	.003
Within Groups	673	86,302.57	128.24		
Total	676	88,068.56			

* $p < .05$

Post Hoc comparisons using the Tukey HSD indicated that the mean score for the *No attendance* condition ($M = 69.50$, $SD = 11.80$) was significantly different from the *Regular attendance with one attempt* condition ($M = 73.90$, $SD = 11.98$), and the *Regular attendance with multiple attempts* condition ($M = 74.88$, $SD = 11.41$). However, the *Irregular attendance* condition did not significantly differ from other conditions. Taken together, these results suggest that learners who attended to self-assessment quizzes of each module at least once and more had higher scores on final exam than others.

Do learner satisfaction and perceived learning differ significantly with respect to self-assessment quiz taking behaviour?

87% of learners answered the course evaluation survey as reported in Table 7.

Table 7: Number of learners who answered the course evaluation survey

	Mean	# of response	
Perceived learning (Over 5)	3.76	591	87%
Course satisfaction level (Over 5)	3.77	591	87%
Number of Suggestions/Comments		123	18%

A one-way between subjects ANOVA was conducted to compare the effect of self-assessment quiz taking behaviour on perceived learning in conditions Irregular attendance, Regular attendance with one attempt, Regular attendance with multiple attempts. Table 8 presents the one-way ANOVA results, which demonstrate the significant effect of self-assessment quiz taking behaviour on perceived learning at the $p < .05$ level for the four conditions [$F(2, 588) = 3.58, p = 0.028$].

Table 8: One-Way Analysis of Variance of Perceived Learning

Source	df	SS	MS	F	p
Between Groups	2	6.20	3.10	3.58	.028
Within Groups	588	508.62	.86		
Total	590	514.82			

* $p < .05$

Post Hoc comparisons using the Tamhane's T2 indicated that the mean score for the *Regular attendance with one attempt* condition ($M = 3.64, SD = 0.90$) was significantly different from the *Regular attendance with multiple attempts* condition ($M = 3.90, SD = 0.85$). The *Irregular attendance* condition ($M = 3.74, SD = 0.99$) did not significantly differ from other conditions. These results suggest that learners who attended to self-assessment quizzes of each module at least once and more perceived a high level of learning than others.

A one-way between subjects ANOVA was conducted to compare the effect of self-assessment quiz taking behaviour on satisfaction in conditions Irregular attendance, Regular attendance with one attempt, Regular attendance with multiple attempts. Table 9 presents the one-way ANOVA results, which demonstrate the significant effect of self-assessment quiz taking behaviour on satisfaction at the $p < .05$ level for the four conditions [$F(2, 588) = 4.40, p = 0.013$].

Table 9: One-Way Analysis of Variance of Satisfaction

Source	df	SS	MS	F	p
Between Groups	2	7.45	3.72	4.40	.013
Within Groups	588	499.14	.84		
Total	590	506.59			

* $p < .05$

Post Hoc comparisons using the Tamhane's T2 indicated that the mean score for the *Regular attendance with one attempt* condition ($M = 3.64, SD = 0.89$) was significantly different from the *Regular attendance with multiple attempts* condition ($M = 3.92, SD = 0.83$). The *Irregular attendance* condition ($M = 3.73, SD = 1.01$) did not significantly differ from other conditions. These results suggest that learners who attend to self-assessment quizzes of each module at least once and more were more satisfied than others.

In course evaluation survey, 18% of learners had some suggestions to improve the course or made some comments about the course. 123 entries in total were made. 11 learners evaluated self-assessment quizzes in their entries. 64% of them advised to increase the number questions in self-assessment quizzes. Learners' direct quotations were shared in Table 10.

Table 10: Learners' direct quotations, which are related to self-assessment quizzes

Learner (L)	Direct Quotation
L1	Self-assessment quizzes improved permanent learning for me
L2	Self-assessment quizzes were helpful
L3	Changing the questions in each attempt resulted in better learning
L4	Number of self-assessment quizzes should be increased
L5, L6	Number of questions should be increased in self-assessment quizzes
L7, L8, L9	Number of True/False questions should be increased in self-assessment quizzes
L10, L11	Increasing number of questions, which aim to promote higher order thinking skills, could increase student engagement

Does self-assessment quiz taking behaviour differ significantly with respect to learner profile?

A Chi-square test of independence was calculated to compare the *self-assessment quiz taking behaviour* and *gender*. No significant relationship was found between these variables ($X^2 = 4.84$, $N = 677$, $p = 0.184$).

A Chi-square test of independence was calculated to compare the *self-assessment quiz taking behaviour* and *their departments*. There wasn't a significant relationship between *self-assessment quiz taking behaviour* and *their department*, ($X^2 = 20.43$, $N = 677$, $p = 0.494$).

Do self-assessment test events differ significantly with respect to learner profile?

6092 test events, which belong to 677 learners, across seven modules on Sakai CLE LMS, were analyzed as represented in Table 11.

Table 11: Descriptive statistics of Test Events

	# of test Events				Test Score Average		
	Female	Male	Total	%	Female	Male	Total
No Attempt	554	690	1244	20.42%			
Attempt Once	1640	1713	3353	55.04%	69.79	68.33	69.04
Multiple attempts	867	628	1495	24.54%	71.77	67.24	69.87
Grand Total	3061	3031	6092	100.00%			

In 20.42% of test events, no attempts were made to take the test as reported in Table 12. In 57.37% of test events, resulted in just one attempt, learners took the test once. In 22.21% of test events, resulted in multiple attempts, learners took the test multiple times.

Table 12: Descriptive statistics of Test Attempt Numbers

Test Attempt	# of Test Event	%
No Attempt	1244	20.42
1	3495	57.37
2	689	11.31
3	225	3.69
4	142	2.33
5	92	1.51
6	61	1.00
7	39	0.64
8	32	0.53
9	24	0.39
10 and above	49	0.80
Total	6092	100.00

An independent-samples t-test was conducted to compare self-assessment test scores in *Attempt Once* and *Multiple Attempts* conditions. There was not a significant difference in the scores for Attempt Once ($M = 69.05$, $SD = 30.47$) and Multiple Attempts ($M = 69.87$, $SD = 29.40$) conditions; $t(2964.68) = 0.89$, $p = 0.371$.

A Chi-square test of independence was performed to examine the relation between the attempt numbers and gender. A significant relationship between attempt number and gender was found ($X^2 = 34.15$, $N = 4848$, $p < .001$). 58% of multiple attempts were made by female learners.

Discussion and conclusion

It is important to understand how self-assessment quizzes affect course achievement in an online course. In this regard, this paper analysed self-assessment quiz taking behaviour of learners in Project Culture online course, which was delivered by one of the biggest and oldest foundation university in western Turkey in fall of 2015. Nikou and Economides (2016) reported that computer and mobile based self-assessment increased science learning motivation and contributed to better learning achievement. Similarly, Belski, (2010) reported that self-assessment could result in better engagement and improve students' course performance as well as Wilson et al. (2011). Similarly, our results suggested that learners who attended to self-assessment quizzes of each module at least once and more, had higher scores on final exam than others. However, our analysis is not strong enough to claim a correlation or cause-and-effect relationship. We can say that there is a relation between taking self-assessment quizzes and final exam scores but we cannot state a direction or a cause-and-effect. There could be two possible situations for the relation we observed. First, learners who had high self-regulation skills took those quizzes. They might be academically more successful students, since they were skilful at self-regulating their learning. So, they had higher scores. Second, learners took higher scores since self-assessment quizzes helped self-regulation of themselves. It was impossible to track overall learning process of the learners with the available log file data. Therefore, we need further research in this topic.

We observed that learners who attended to self-assessment quizzes of each module at least once and more perceived a higher level of learning than others. In addition, they were more satisfied than others. These findings are parallel to other findings in the field of self-assessment study; according to researchers (Brown & Harris, 2013; McMillan & Hearn, 2008; Oscarson, 1989; Panadero & Alonso-Tapia, 2013) correctly implemented self-assessment is by itself a process that promotes learning.

In this study 57.37% of test events resulted in just one attempt, and no statistically significant difference was found in the test scores for Attempt Once and Multiple Attempts. This means that although attending self-assessment quizzes regularly implies a difference on final exam scores, number of attempts to those quizzes does not have an effect on the scores. Therefore, allowing two or three attempts for each self-assessment quiz will probably be sufficient to achieve the objectives of the self-assessment. It seems like creating an extra system load by allowing unlimited attempts is not necessary.

No significant relationship was found between self-assessment quiz taking behaviour and learner profile in terms of their gender and department. However, a significant relationship between *attempt number* and *gender* was observed. 58% of multiple attempts were made by female learners. In her speech on TED Talks, Saujani (2016), the founder of Girls Who Code, gives some statistics and states that females are more perfectionist than males since we are raising our girls to be perfect while we are raising our boys to be brave. Her detection and intuition are true for Turkish culture, as well as American culture. In this respect, female learners might have a tendency to reach to the highest score, in other words to reach the perfect score. So, the reason of higher percentage rate of multiple attempts in favour of female learners could be this tendency. However, this issue requires further research.

These findings could be references for those who plan to design and develop self-assessment activities in online courses. In addition, the results of our analysis could be helpful to policy decision makers in order to support their educational development. To improve external validity, it is planned to replicate the analyses in next semesters.

Resources

1. Adams, T. L. (1998) Alternative assessment in elementary school mathematics. *Childhood Education*, 74(4), 220–224.
2. Andrade, H., & Boulay, B. (2003) Gender and the role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research*, 97(1), 21–34.
3. Bälter, O., Enström, E., & Klingenberg, B. (2013). The effect of short formative diagnostic web quizzes with minimal feedback. *Computers & Education*, 60(1), 234–242.
<http://dx.doi.org/10.1016/j.compedu.2012.08.014>.
4. Belski, I. (2010). The Impact of Self-Assessment and Reflection on Student Learning Outcomes. *Proceedings of the 2010 AaeE Conference, Sydney*, 216-221. Retrieved from https://www.researchgate.net/publication/228460756_The_Impact_of_Self-Assessment_and_Reflection_on_Student_Learning_Outcomes
5. Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151-167. doi: 10.1080/713695728.
6. Boud, D., & Brew, A. (1995). Developing a typology for learner self assessment practices. *Research and development in Higher Education*, 18, 130–135.
7. Boud, D., & Falchikov, N. (1989). Quantitative studies of student self-assessment in higher-education: A critical analysis of findings. *Higher Education*, 18(5), 529-549.
doi:10.1007/BF00138746
8. Bourke, R. (2010). *The chameleonic learner. Learning and self-assessment in context*. Wellington: New Zealand Council for Educational Research.

9. Brown, G. T., & Harris, L. R. (2013). Student self-assessment. In J. McMillan (Ed.), *The SAGE handbook of research on classroom assessment* (pp. 367-393). CA: SAGE.
10. Carter, J., Ala-Mutka, K., Fuller, U., Dick, M., English, J., Fone, W., et al. (2003). How shall we assess this? *SIGCSE Bulletin*, 35, 107–123.
11. Claxton, G. (1995). What Kind of Learning does Self-assessment drive? Developing a ‘Nose’ for Quality: Comments on Klenowski. *Assessment in Education: Principles, Policy and Practice*, 2(3), 339-343.
12. Ćukušić, M., Garača, Z., & Jadrić, M. (2014). Online self-assessment and students’ success in higher education institutions. *Computers & Education*, 72(2), 100–109.
13. Deutsch, T., Herrmann, K., Frese, T., & Sandholzer, H. (2012). Implementing computer-based assessment – a web-based mock examination changes attitudes. *Computers & Education*, 58, 1068–1075.
14. Dochy, F., & McDowell, L. (1997) Introduction: assessment as a tool for learning. *Studies in Educational Evaluation*, 23(4), 279–298.
15. Falchikov, N. & Boud, D. (1989) Student self-assessment in higher education: a meta-analysis. *Review of Educational Research*, 59(4), 395–430.
16. Frederiksen, J., & Collins, A. (1989) A systems approach to educational testing. *Educational Researcher*, 18(9), 27–32.
17. Gagne, R. M., Briggs, L. S., & Wager, W. W. (1988). *Principle of Instructional Design*. Holt: Rinehart and Winston.
18. Gale, J. (1984). Overview: Self-assessment and Self-remediation Strategies. In E. S. Henderson & M. B. Henderson (Eds.), *Independent Learning in Higher Education*. NJ: Educational Technology Publications.
19. Garcia, J. A., & Floyd, C. E. (1999) Using single system design for student self-assessment: a method for enhancing practice and integrating curriculum. *Journal of Social Work Education*, 35(3), 451–461.
20. Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: a review of the literature. *Computers & Education*, 57(4), 2333–2351. doi:10.1016/j.compedu.2011.06.004
21. Goodrich, H. (1996). *Student self-assessment: at the intersection of metacognition and authentic assessment*. Unpublished doctoral dissertation, Harvard University.
22. Gretes, J. A., & Green, M. (2000). Improving undergraduate learning with computer-assisted assessment. *Journal of Research on Computing in Education*, 33(1), 46–54.
23. Heidi A., & Du, Y. (2007) Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, 32(2), 159-181, doi:10.1080/02602930600801928
24. Hanrahan, S. J., & Isaacs, G. (2001) Assessing self- and peer-assessment: the students’ views. *Higher Educational Research and Development*, 20(1), 53–70.
25. Horner, S., & Shwery, C. (2002) Becoming an engaged, self-regulated reader. *Theory into Practice*, 41(2), 102–109.

26. Ibabe, I., & Jauregizar, J. (2009). Online self-assessment with feedback and metacognitive knowledge. *Higher Education*, 59, 243–258.
27. Kulik, J. A., Kulik, C.-L., & Bangert, R. L. (1984). Effects of practice on aptitude and achievement test scores. *American Educational Research Journal*, 21(2), 435–447.
28. Laurillard, D. (1993). *Rethinking University Teaching: a framework for the effective use of educational technology*. NY: Routledge.
29. McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational HORIZONS*, 87(1), 40-49.
30. Nikou, S. A., & Economides, A. A. (2016). The impact of paper-based, computer-based and mobile-based self-assessment on students' science motivation and achievement. *Computers in Human Behavior*, 55, 1241–1248. doi:10.1016/j.chb.2015.09.025
31. Oscarson, M. (Ed.). (1978). *Approaches to self assessment in foreign language learning*. London: Pergamon Press.
32. Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, 6(1), 1-13. doi:10.1177/026553228900600103
33. Panadero, E. (2011). *Instructional help for self-assessment and self-regulation: Evaluation of the efficacy of self-assessment scripts vs. rubrics*. Ph.D. dissertation. Spain: Department of Educational and Developmental psychology, Universidad Autónoma de Madrid.
34. Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: Theoretical and Practical Connotations. When it Happens, How is it Acquired and what to do to Develop it. *Electronic Journal of Research in Educational Psychology*, 11(2), 551-576. doi:10.14204/ejrep.30.12200
35. Paris, S. G., & Paris, A. H. (2001) Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89–101.
36. Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in Education: Theory, Research, and Applications* (2nd ed.). N.J.: Merrill/Prentice Hall.
37. Rolheiser, C., & Ross, J. A. (2013). *Student Self-Evaluation: What Research Says and What Practice Shows*. Retrieved from http://moodle.manistee.org/pluginfile.php/59439/course/section/16807/STUDENT_SELF-EVALUATION_WHAT_RESEARCH_SAYS_AND_WHAT_PRACTICE_SHOWS.pdf
38. Rowntree, D. (1991). *Teaching Through Self-instruction: How to Develop Open Learning Materials*. NY: Nichols Publishing.
39. Schunk, D. H. (2004). *Learning Theories: An Educational Perspective*. N.J.: Merrill Prentice/Hall.
40. Saujani, R. (2016, February). Teach girls bravery, not perfection. [Video file]. Retrieved from https://www.ted.com/talks/reshma_saujani_teach_girls_bravery_not_perfection#t-6027
41. Snooks, M. K. (2004). Using practice tests on a regular basis to improve student learning. *New Directions for Teaching and Learning*, 100, 109–113.
42. Stiggins, R. J. (2001). *Student-involved classroom assessment* (3rd ed.). NJ: Merrill/Prentice-Hall.
43. Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston: Pearson.

44. Taylor, J. A. (1998). Self Test: a flexible self assessment package for distance and other learners. *Computers & Education*, 31(3), 319–328. [http://dx.doi.org/10.1016/S0360-1315\(98\)00044-X](http://dx.doi.org/10.1016/S0360-1315(98)00044-X)
45. Taras, M. (2005). Assessment – summative and formative – some theoretical reflections. *British Journal of Educational Studies*, 53, 466–478.
46. Wiggins, G. (1998). *Educative assessment: designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass.
47. Wilson, K., Boyd, C., Chen, L., & Jamal, S. (2011). Improving student performance in a first year geography course: examining the importance of computer-assisted formative assessment. *Computers & Education*, 57(2), 1493–1500. <http://dx.doi.org/10.1016/j.compedu.2011.02.011>.
48. Zhang, D., Zhao, J. L., Zhou, L., & Nunamaker Jr. J. F. (2004). Can e-learning replace classroom learning? *Communications of the ACM*, 47, 75–79.
49. Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2). Retrieved from https://www.researchgate.net/publication/237065878_Becoming_a_Self-Regulated_Learner_An_Overview