Ampirik Araştırma / Original Empirical Research

www.yuksekogretim.org

TÜBA-HER Yükseköğretim Dergisi
TÜBA Higher Education Research/Review

Evaluating the Accuracy of Actual Workload Questionnaire with Ordinal Response Category: Bootstrap Sampling for Reliability

Sıralı Yanıt Kategorili Gerçek İş Yükü Anketinin Doğruluğunun Değerlendirilmesi: Güvenilirlik için Bir Bootstrap Örnekleme

Selcen Yüksel[1], Pervin Demir[1], Şemsettin Tabur[2], Ankara Yıldırım Beyazıt University Quality Board[3]

[1] Department of Biostatistics, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Ankara, Türkiye
[2] Department of Translation and Interpretation, Ankara Yıldırım Beyazıt University, Faculty of Humanities and Social Sciences, Ankara, Türkiye
[3] Rectorate, Ankara Yıldırım Beyazıt University, Ankara, Türkiye

Özet

Bologna Deklarasyonu (19 Haziran 1999), üniversite derecelerinin müfredatlarının dönüştürülmesini sağlayan Avrupa Yüksek Öğretim Alanı'nın (AYÖA) uygulamalarını motive etti. Türkiye'de yükseköğretimde kalite güvencesi ağırlıklı olarak Türkiye'nin 2001 yılında Bologna Süreci'ne katılmasından sonra gelişmeye başlamıştır. Bologna reformları ile ilgili çaba ve faaliyetler özellikle 2005 yılından sonra hız kazanmıştır. AYÖA çerçevesinde, öğrenci iş yükü, belirli bir dersi geçmek için gereken toplam sınıf içi ve bağımsız yapılandırılmamış çalışmanın sayısal tanımlayıcı bir değeri olan Avrupa Kredi Transfer Sistemi (AKTS) aracılığıyla belirlenir. AKTS'nin hesaplanmasında birincil girdi olan öğrencinin fiili iş yükü, genellikle öğretim elemanının ders dışı zamanlarda öğrenciden beklentileri dikkate alınarak hesaplanır. Bazı üniversitelerde öğretim elemanları aldıkları dersle ilgili fiili iş yükü anketini yapmak ve sonuçları bölümleri ile paylaşmakla yükümlüdürler. Bazı öğretim elemanları için böyle bir yaptırım olsa da çoğu üniversitede öğrencilerin gerçek iş yükü doğrudan öğrencilere sorulmadan tahmin edilmektedir. Bu çalışmada, öğrencilerin fiili iş yüklerini hesaplamak için yapılandırılmış bir anket tasarlanması ve bu anketten elde edilen sonucun güvenilirliğinin istatistiksel olarak kanıtlanması amaçlanmıştır. Bu amaçla 3 ders için sıralı yanıt kategorisine sahip gerçek iş yükü anketi öğrencilere uygulanmış, aynı anketin açık uçlu versiyonu yine aynı öğrencilere yüz yüze sorulmuş ve sonuçlar karşılaştırılmıştır. Ayrıca, bu karşılaştırmanın doğruluğunu değerlendirmek ve sonuçların popülasyona genellenebilirliğini sağlamak için, anketlerden alınan yanıtların ampirik dağılımları dikkate alınarak, bootstrap örnekleme yöntemi kullanılmıştır.

Anahtar Kelimeler: AKTS, Gerçek İş Yükü Anketi, Bootstrap Örnekleme, Güvenirlik

Abstract

The Bologna Declaration (19 June 1999) motivated the implementation of the European Higher Education Area (EHEA), which led to the transformation of the curricula of the university degrees. In Türkiye, quality assurance in higher education was focused mainly after the participation of Türkiye in Bologna Process in 2001. However, efforts and activities related to Bologna reforms gained momentum particularly after 2005. Within the EHEA framework, the student workload is set through the European Credit Transfer and Accumulation System (ECTS), which is a numerical descriptive value of the total in-class and independent unstructured work required to pass a particular course. The actual workload of the student, which is the primary input in the calculation of ECTS, is generally calculated by taking into account the lecturer's expectations from the student during extracurricular times. In some universities, lecturers are obliged to conduct the actual workload questionnaire related to the course they are taking and share the results with their departments. Although there is such a sanction for some lecturers, the actual workload of students is estimated in most universities without asking students directly. In this study, it is aimed to design a structured questionnaire to calculate the actual workload of students and to statistically prove the reliability of the result obtained from this questionnaire. For this purpose, actual workload survey results with ordinal response categories for 3 courses were obtained and the open-ended version of the same questionnaire was asked to same students again face to face and the results were compared. In addition, to evaluate the accuracy of this comparison and the generalizability of the results to the population was questioned using the bootstrap sampling method, taking into account the empirical distributions of responses got from questionnaires.

Keywords: ECTS, Actual Workload Questionnaire, Bootstrap Sampling, Reliability

İletişim / Correspondence:
Prof. Dr. Selcen Yüksel
Department of Biostatistics, Ankara Yıldırım Beyazıt University, Faculty of Medicine, Ankara / Türkiye
e-mail: selcenyuksel@aybu.edu.tr

Yükseköğretim Dergisi / TÜBA Higher Education Research/Review (TÜBA-HER), 15(ÖS), 1–6. © 2025 TÜBA
Geliş tarihi / Received: Ağustos / August 27, 2024; **Kabul tarihi / Accepted:** Ocak / January 1, 2024
Bu makalenin atıf künyesi / How to cite this article: Yüksel, S., Demir, P., Tabur, Ş. & Ankara Yıldırım Beyazıt University Quality Board. (2025). Evaluating the accuracy of actual workload questionnaire with ordinal response category: Bootstrap sampling for reliability. *Yükseköğretim Dergisi, 15*(ÖS), 11–16. https://doi.org/10.53478/yuksekogretim.1539693

ORCID: S. Yüksel: 0000-0002-8994-8660; P. Demir: 0000-0002-6652-0290; Ş. Tabur: 0000-0002-8648-8558; Ankara Yıldırım Beyazıt University Quality Board: 0009-0007-5615-3144

The Bologna Declaration of 19 June 1999 aimed at implementing European Higher Education Area (EHEA), which has further led to the transformation of curricula of the university degrees (Egea et al., 2022). In Türkiye, quality assurance processes in higher education started to gain significance mainly after Türkiye's participation in Bologna Process in 2001. However, activities related to the Bologna Process gained momentum particularly after 2005. Within the EHEA framework, the student workload is set through the European Credit Transfer and Accumulation System (ECTS), which is a numerical descriptive value of the total in-class and independent unstructured work required to pass a particular course and it is used to determine the number of academic credits that should be assigned to a specific course or activity. ECTS system has also been implemented to compare and transfer academic credits between higher education institutions in the European Union and other participating countries. (Souto-Iglesias et al., 2018). Universities and colleges use ECTS credits to make it easier for students to compare and transfer credits between institutions and countries, which, in turn, helps students with their academic mobility. Therefore, ECTS is an important tool for accrediting institutions and degree programs, for it ensures that academic credits earned at one institution are comparable to those earned at another institution. In the accreditation process, student workload is often used as a measure of the rigor and quality of the education provided by an institution or program. Accrediting bodies often look at the number of hours that students are expected to spend on coursework and other activities, as well as the types of assignments and assessments they are required to complete. By examining student workload, accrediting bodies can gain insight into the level of engagement and rigor of the program and the institution.

The Higher Education Quality Council of Türkiye (THEQC) is an independent organization responsible for overseeing and coordinating the quality of higher education in Türkiye. Two main functions of THEQC are to accredit institutions of higher education and provide them with guidance and support in terms of improving the quality of their education. THEQC, makes external institutional evaluations by visiting its institutions and gives accreditation to successful universities that meet quality criteria through The Institutional Accreditation Program (IAP) since 2016. IAP is an external evaluation method that enables the evaluation of quality assurance, learning and teaching, research and development, social contribution, and governance system processes of higher education institutions in accordance with the plan-do-check-act cycle (THEQC, 2022a), IAP is conducted by the evaluation teams assigned by THEQC in compliance with the Institutional External Evaluation and Accreditation Criteria and the Institutional External Evaluation and Accreditation Guide.

Student workload questionnaire result is major evidence for IAP in accreditation process to measure student engagement, rigor, and overall satisfaction with the course or program, to assess the quality and integrity of the institution or program. It is typically used to gather information about how much time and effort students spend on their coursework, as well as to gather feedback on the quality of the course or program. These questionnaires are often conducted by professors or instructors, academic advisors, or program directors, who use the results to evaluate the quality of the course or program and make necessary adjustments for improvement. As mentioned above, student workload is an important issue in the accreditation processes for institutions and programs.

In the last two years, 23 institutions have been accredited by THEQC with the IAP system (THEQC, 2022b). While 13 of them received full accreditation, 10 received conditional accreditation. When the reports of 13 institutions were assessed, applying student workload questionnaires and the evaluating of the results systematically were marked as areas open for improvement in the reports of 5 institutions. In 9 out of 10 conditionally accredited institutions, the evidence regarding the actual workload was found to be insufficient. In order for this evidence to be satisfactory, student workload questionnaires, measuring the number of hours a student is expected to allocate for a course with open-ended questions, should be conducted by the course or program instructor for each course in every semester. Due to the workload of course instructors, it is essential to develop mechanisms that assist them with this process. Most of the student information systems used by Turkish universities are available for administering multiple choice questionnaires to all students simultaneously. Even if the calculation of the actual workload with open-ended questions is the gold standard, applying multiple choice questionnaires to measure it will facilitate the achievement of real workloads by providing simultaneous access to all the courses taken by all students on the student information system. In designing this study, we aimed to check the level of correspondence between the answers taken from 3 short multiple-choice questions and the open-ended answers received using the student information system at the end of each term to determine the actual workload for any course. To this end, we first applied the multiple-choice student workload questionnaire to 4 courses on the student information system. Then, the open-ended real workload questionnaire was applied to the participants of the same courses after their final exam. We evaluated the concordance of the answers received in both methods. The bootstrap method was used to measure the uncertainty in the concordance and the calculated statistics. In doing so, we assessed the uncertainty of a statistic by simulating the sampling process multiple times, using a different random sample from the original dataset each time.

Method

The study sample consisted of the responses received from the students who visited the courses IMT219 (Cultural History of the United States) and IMT322 (English and American Short Stories) in the Faculty of Humanities and Social Sciences, and SBF102 (Nutrition and Physiotherapy) in the Faculty of Health Sciences at Ankara Yildirim Beyazit University. The total sample size for each lecture was 22, 12, 61, and 52, respectively.

A student's workload for a course includes both in-class and out-of-class activities. The lecture and exam duration can be determined by the instructor for in-class activities; however, the time spent for out-of-class activities should be questioned directly from the students. The out-of-class activities refer to the time that students spend on independent learning and studying for exams. Thus, the following three questions were directed at students in order to evaluate the amount of time spent for out-of-class activities and to calculate their overall workload:

Q1. How many hours per week did you spend on average outside of class hours in this course?
Q2. How many hours did you spend on average for the midterm exam preparation in this course?
Q3. How many hours did you spend on average for the final exam preparation in this course?

Then, the students' workload and the ECTS were calculated by the equations given below.

Total hours of instruction= the scheduled hours (the semester lasts approximately 14 weeks*2 hours/week) + the average time for midterm and final exams (2 hours)

Total hours out-of-class independent study= (the semester lasts approximately 14 weeks *response to Q1) + response to Q2 + response to Q3.

The students' workload = total hours of instruction + hours out-of-class independent study hours

As one ECTS credit represents a workload of 25-30 hours (Directorate-General for Education, Youth, Sport, and Culture, 2015), the ECTS of a course was calculated by dividing the total workload by 30.

The ECTS of course = The students' workload / 30

The questionnaire was administered in two different ways:

1- The questions were managed with ordinal response categories (0, 1-3, 4-6, 7-9, 10-12, and 913 hours) for the selected four courses on student affairs information system (SAIS). The participants were asked to mark the option which is nearest to the real value. The middle value of the category specified in the workload calculation was defined as "*SAIS*". For example, suppose that the answer of a student on "How many hours per week did you spend on average for this course outside of class hours?" question is "*4-6*". We defined this student's answer as "*5*" in our calculations. If the student answered a question in 13 hours or more, we fixed the student's answer at 14 numerically.
2- The open-ended version of the same questionnaire was asked in person to the same students again after final exam. In this case, the hour directly specified by the student was taken in the workload calculation and defined as "*Real*".

Due to personal data privacy of students, the answers given in SAIS were anonymous and thus could not be directly matched with open-ended answers received in person. Therefore, in order to evaluate the concordance between the categorical results obtained in the 1st stage and the open-ended answers obtained in the 2nd stage, we found the equivalent of the open-ended answers given in the 2nd stage according to the categories given in the 1st stage, and we defined new categories as "*Converted*". Suppose a student replied the open-ended "How many hours did you spend on average the preparation for the midterm exam for this course?" with 6. Accordingly, we noted that the "*Real*" value of this student is 6. Since if the student's answer is 6, this answer corresponds to the 3rd category in SAIS and the middle point of this category is 5, we got the "*Converted*" value of the student as 5.

Statistical Analysis

The median of the students' workload and the ECTS were calculated for each course. The concordance between the ECTS values obtained from the open-ended version ("*Real*") and over their categorical equivalents ("*Converted*") was evaluated by intraclass correlation coefficient (ICC, two-way mixed model, agreement, single). The bootstrap samples with replacement were created instead of collecting the large data to build up the sampling distribution. The bootstrap statistic is a nonparametric resampling method which makes no distributional assumptions and used for estimation.

The basic steps of the bootstrap method used in this study are as follows:

1- We took a random sample with replacement from the original dataset. This sample is called a bootstrap sample.
2- We calculated the statistics of interest for the bootstrap sample. The main two statistics are median of ECTS for each lecture and ICC of "*Converted*" and "*Real*" values.
3- We repeated steps 1 and 2 10000 times to generate a large number of bootstrap samples and corresponding statistics.

4- The mean, bias, and the root mean square residuals (RMSR) of the statistic values based on the 10000 ordinary nonparametric bootstrap replicates was determined given formulas, respectively.

$$MEAN=\frac{\sum_{i:1}^{R} t_i}{R} \qquad BIAS=\frac{\sum_{i:1}^{n}(t_i-t)}{R} \qquad RMSR=\sqrt{\frac{\sum_{i:1}^{R}(t_i-t)^2}{R}}$$

t_i: the estimated statistic value of sample based on the ith bootstrap replicate.
R: the number of nonparametric bootstrap replicates
t: the statistic value from the original study sample

Accuracy of statistical significance of median and ICC estimates were quantified by the 2.5th and 97.5th percentiles within the sorted distribution of the estimated parameters as the lower and upper limits of the bootstrap confidence intervals (Efron, 1987).

The "*boot*" (Canty & Ripley, 2022) and "*irr*" (Gamer et al., 2022) packages in RStudio were used for the estimation of the median and ICC values, respectively. "*boot*" and "*irr*" are R libraries that generates R bootstrap replicates of a statistic applied to data and calculates interrater reliability and agreement, respectively.The mean of estimations was obtained from the replicates.

Findings

The median ECTS value of IMT218, IMT322, SBF102-Nutrition, and SBF102-Physiotherapy lectures were presented in Table 1. The ECTS median values obtained from the open-ended version and ordinal response categories version were similar for these lectures.

Table 1
The median ECTS values for each lecture

	IMT218	IMT322	SBF102 - Nutrition	SBF102 - Physiotherapy
Real	2.42	3.48	2.13	1.97
Converted	2.52	3.62	2.37	2.27
SAIS	2.27	3.67	2.37	2.42

The ICC between the Real and Converted ECTS values were obtained from each bootstrap samples, and the mean of the 10000 ICC values were calculated. Table 2 below shows the ICC values with the percentile 95% CIs from the original study sample and the bootstrap samples. The agreements between the Real and Converted ECTS values from the bootstrap samples were obtained as excellent (0.97), good (0.87) and moderate (0.74 and 0.71), respectively.

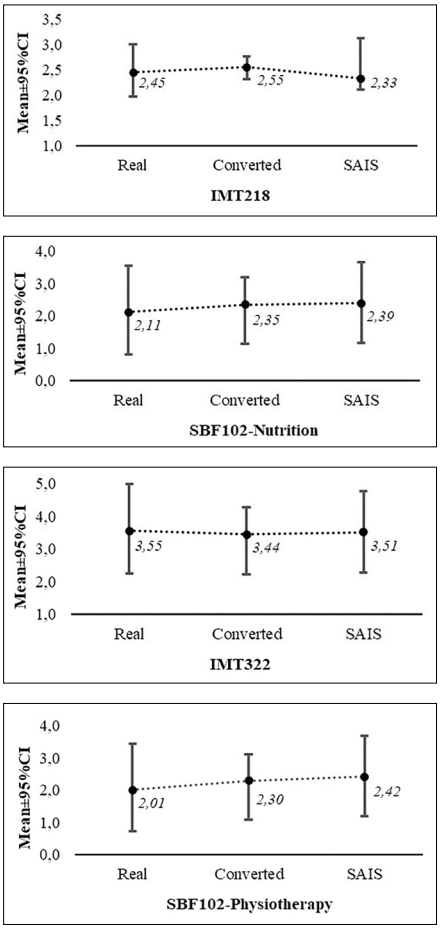
Table 2
The ICC between the Real and Converted ECTS values from the original study sample and the bootstrap samples

	ICC (95% CI) from the original sample	ICC (95% CI) from the bootstrap samples
IMT218	0.97 (0.94 – 0.99)	0.97 (0.94 – 0.99)
IMT322	0.89 (0.68 – 0.97)	0.87 (0.71 – 0.97)
SBF102 - Nutrition	0.69 (0.53 – 0.80)	0.74 (0.59 – 0.94)
SBF102 - Physiotherapy	0.66 (0.47 – 0.79)	0.71 (0.57 – 0.87)

ICC estimates were classified as following rules: 0.50-0.75 indicate moderate agreement, 0.76- 0.90 indicate good agreement, and >0.90 indicate excellent agreement (Koo & Li, 2016).

The mean of the ECTS estimates from the bootstrap samples with 10000 replacements were reported with the 95%CIs at Figure 1 (a-d). CIs for ECTS estimates obtained from "Real" data included lower and upper bounds of the CIs obtained from "Converted" data. Similarly, the ranges of ECTS estimation values calculated from "Real" and "SAIS" data for four lectures overlap.

Figure 1 (a-d)
The mean and percentile 95% CIs estimated median and percentile 95% CIs of ECTS median values



Discussion and Conclusion

In this study, it was investigated whether the student workload could be obtained reliably from the student information system. For this purpose, it was attempted to reach the result of the population with bootstrap samples by using hypothetic data sets for 4 courses.

In its internal evaluation and institutional accreditation procedures, THEQC has clearly set out its criteria regarding the necessity of student workload to be a well-designed, publicized and systematically-implemented process conducted by universities. Higher education institutions are therefore required to carry out monitoring and improvement studies on student workload. According to the 2021 quality assurance status report published by THEQC (THEQC, 2022c), institutions are generally successful in determining and implementing defined processes concerning the design and approval of study programs. However, mechanisms for monitoring and ensuring continuous improvement of these processes are available in only of the 42% of institutions. It is noteworthy that one of the most important problems observed in institutions under this criterion is the design processes based on student workload. According to the report, while monitoring and improvement of student workload based on feedback is carried out in only 3 institutions, mechanisms for monitoring the workload in other institutions are not mature yet. In 90% of the universities conditionally accredited by IAP, deficiencies were detected in the evaluation process of the student workload. At this point, it is clear that it is necessary to develop efficient, standardized and easy-to-apply mechanisms in institutions.

In the literature, open-ended questions are asked to the students in order to determine the extracurricular work time that should be known in the calculation of the student workload. It is important to use fast and easy-to-use tools in order to carry out this calculation easily and to make the evaluation process practical in all the other courses offered in a university. In calculating the student workload, the use of student information system, which provides access to all students simultaneously for all courses, can be considered as the fastest and feasible? method. However, it is not always possible to make open-ended inquiries in these systems. Even with open-ended querying, systems often do not allow to get specific calculations from open-ended results. In our study, the reliability of the sequential questionnaire defined on the student information system, which is considered to be the fastest method in calculating the student workload for 4 courses, on defining the student workload was examined. For this purpose, in the 10000 bootstrap samples obtained for 4 courses, medium and high agreement was found between *Real* and *Converted* results (0.71, 0.74, 0.87 and 0.97) (Table 2). This result proves that multiple-choice questions can be used to calculate student workload for courses instead of open-ended questions. When the confidence intervals obtained from the bootstrap simulation were examined, it was observed that the lower limits of all confidence intervals were above 50% (Table 2). This is a proof for the significance of the ICC mean values. It is seen that the confidence intervals for the ECTS estimates obtained from the *Real* data include the confidence intervals obtained from the *Converted* data. Similarly, the intersection of the ECTS confidence intervals determined by the *Real* values and the ECTS confidence intervals determined by the *SAIS* values demonstrated that the queries made with two different methods gave similar results (Fig. 1). Also, median ECTS values got from bootstrap samples gives near values for all four lecture (Table 1). It means that the ECTS median values obtained from the open-ended version and ordinal response categories version were similar. The strength of our study is that the bootstrap method was used for population generalization of the obtained estimates, considering that the number of students in the departments was low and the ECTS distribution calculated over the questions directed to the workload calculation was not known. Indeed, including four lectures is a limitation for this study. For this reason, it is important to support the results by making similar evaluations for different courses in different faculties and departments or design specific simulation process to answer this question taking various simulation conditions.

As a result, in calculating the student workload for a lecture, calculations can be made using the answers obtained with the multiple-choice questionnaire in student information systems that do not facilitate questioning with an open-ended questionnaire. In addition, the rapid calculation of the workload through the student information system in each academic year will ensure the up-to-dateness in the calculation of the ECTS of relevant course.

References

Canty, A., & Ripley, B. (2022). *Boot: Bootstrap R (S-Plus) functions*. R package version 1.3-28.1. https://brieger.esalq.usp.br/CRAN/web/packages/boot/boot.pdf

Directorate-General for Education, Youth, Sport, and Culture (European Commission). (2015). ECTS Users' Guides. *Publication of European Union*. ISBN:978-91-43562-1. https://doi.org/10.2766/87592.

Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, *82*(397), 171-185. https://doi.org/10.1080/01621459.1987.10478410

Egea, G., Rodríguez-Lizana, A., Pérez-Urrestarazu, L., Pérez-Ruiz, M., Rallo, P., & Suárez, M.P. (2022). Assessment of actual workload and student performance in the agricultural engineering final degree project in a Spanish higher education context. *Education Sciences*, *12*(6), 418. https://doi.org/10.3390/educsci12060418

Gamer, M., Lemon, J., Fellows, I., & Singh, P (2022). irr: Various coefficients of interrater reliability and agreement. R package version 0.84.1. https://cran.r-project.org/web/packages/irr/irr.pdf

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155-163. https://doi.org/10.1016/j.jcm.2016.02.012

Souto-Iglesias, A., & Baeza_Romero, M.T. (2018). A probabilistic approach to student workload: empirical distributions and ECTS. *Higher Education*, *76*(2), 1007–1025. https://doi.org/10.1007/s10734-018-0244-3

Turkish Higher Education Quality Council. (2022a, January). *What is institutional accreditation program?* https://yokak.gov.tr/degerlendirme-sureci/kurumsal-akreditasyon-programi-nedir

Turkish Higher Education Quality Council. (2022b, January). *Accredited institutions and reports*. https://yokak.gov.tr/raporlar/akredite-olan-kurumlar

Turkish Higher Education Quality Council. (2022c, January). *The 2021 higher education evaluation and quality assurance status report*. ISBN:978-975-7912-62-0, https://yokak.gov.tr/Common/Docs/Site_Activity_Reports/DurumRaporu2021.pdf