Ampirik Araştırma / Original Empirical Research

www.yuksekogretim.org

TÜBA-HER Yükseköğretim Dergisi
TÜBA Higher Education Research/Review

Consistency Analysis Inter-Team and Inter-Year in Program Accreditation: TURAK Example from Türkiye

Program Akreditasyonunda Takımlar ve Yıllar Arası Tutarlılık Analizi: Türkiye'den TURAK Örneği

Abdulcelil Çakıcı

Mersin University, Tourism Faculty, Department of Tourism Management, Mersin, Türkiye

Abstract

Although it is relatively easy to ensure consistency within the team, it is very difficult to ensure consistency between teams and across years in accreditation assessments. In the study, an analysis was conducted based on the reports provided by evaluation teams to the Tourism Education Evaluation and Accreditation Board (TURAK) in Türkiye. A numeric rubric system-"1: Deficit, 2: Inadequate, 3: Acceptable, 4: Good, 5: Very good"- was used for fifty standards under nine dimensions. A written rule was established stating that the difference in scores should not exceed 2 for each item. The dimension average is calculated by summing the average scores of all items in the dimension and dividing by the number of items. A dataset was created using reports from the teams, covering the years 2021, 2022, and 2023.An analysis of the nine dimensions revealed that there were not any significant differences in team evaluations by year, university type, or evaluation type. This indicates that the teams made similar evaluations across years, universities, and evaluation type. In other words, this finding implies that consistency was achieved between teams across years, university type and evaluation type.

Keywords: Program Accreditation, Consistency, Inter-Team Consistency, Inter-Year Consistency

Özet

Akreditasyon değerlendirme takım içi tutarlılığı sağlamak nispeten kolay olsa da,takımlar arası ve yıllar arası tutarlılığı sağlamak oldukça zordur. Çalışmada, Türkiye'de Turizm Eğitimi Değerlendirme ve Akreditasyon Kurulu'na (TURAK) değerlendirme takımları tarafından verilen raporlar baz alınarak bir analiz gerçekleştirilmiştir. Dokuz boyut altında elli standart için "1: Eksik, 2: Yetersiz, 3: Kabul edilebilir, 4: İyi, 5: Çok iyi" şeklinde nümerik bir rubrik sistem kullanılmıştır. Her bir madde için puan farkının 2'yi geçmemesi gerektiği yazılı bir kural haline getirilmiştir. Bir boyutun ortalaması, boyuttaki tüm maddelerin ortalama puanlarının toplanıp madde sayısına bölünmesiyle elde edilir. Takımlardan gelen raporlar dikkate alınarak 2021, 2022 ve 2023 yıllarına ait verileri kapsayan bir veri seti oluşturulmuştur. Dokuz boyut dikkate alınarak yapılan analizler; takımların yıllara, üniversite türüne ve değerlendirme türüne göre değerlendirmeleri arasında anlamlı bir fark olmadığını ortaya koymuştur. Bu bulgu, takımların yıllara, üniversitelere ve değerlendirme türüne göre benzer değerlendirmeler yaptığını ortaya koymaktadır. Bir başka deyişle, takımlar arasında yıllara, üniversite türüne ve değerlendirme türüne göre tutarlılık sağlandığı söylenebilir.

Anahtar Kelimeler: Program Akreditasyonu, Tutarlılık, Takımlar Arası Tutarlılık,Yıllar Arası Tutarlılık

Accreditation, defined as "the state of being reliable and believable" (Kılıçaslan, 2020; Semerci, 2017), is accepted as a quality indicator (Kumar et al., 2020). Therefore, it has been witnessed that developing national accreditation systems to ensure the standards in higher education has gained significant attention in European countries, particularly after Bologna Declaration (Prøitz et al., 2004; Sin et al., 2017). Accreditation in higher education can also be viewed as a system aimed evaluating whether an institution or program meets certain performance criteria and sharing the result with the public (Kılıçaslan, 2020).

There are two types of accreditation in higher education: institutional accreditation and program accreditation. While institutional accreditation is carried out within the scope of evaluating national standards in a higher education institution, program accreditation is handled to evaluate the criteria that the program must meet (Arnanz & Kaewnuch, 2019). In Türkiye, institutional accreditation is managed by the Turkish Higher Education Quality Council (THEQC), while program accreditations are carried out by accreditation agents authorized by THEQC (THEQC Regulation, Art.6/1/e).

İletişim / Correspondence:
Prof. Dr. Abdulcelil Çakıcı
Mersin University, Tourism Faculty, Department of Tourism Management, Mersin / Türkiye
e-mail: celilcakici@mersin.edu.tr

Yükseköğretim Dergisi / TÜBA Higher Education Research/Review (TÜBA-HER), 15(ÖS), 103-114. © 2025 TÜBA
Geliş tarihi / Received: Eylül / September 3, 2024; **Kabul tarihi / Accepted:** Şubat / February 5, 2025
Bu makalenin atıf künyesi / How to cite this article: Çakıcı, A. (2025). Consistency analysis inter-team and inter-year in program accreditation: TURAK example from Türkiye. *Yükseköğretim Dergisi, 15*(ÖS), 113-124.. https://doi.org/10.53478/yuksekogretim.1542751

ORCID: A. Çakıcı: 0000-0002-9192-1969

Generally, similar processes are followed in both institutional and program accreditations. In program accreditation, the program seeking accreditation first applies to the relevant accreditation board. If the program's application is accepted, it prepares a Program Self-Assessment Report (PSER) containing evidence-based explanations. This report is then forwarded to the relevant accreditation agent. Then the relevant accreditation body constitutes an evaluation team consisting of field experts who have received evaluator training. The evaluation team is then asked to review the PSER, conduct a site visit, and present orally their evidence-based findings to University's rector and the officials invited by rector, in a meeting named as Exit Statement. Universities may object to the draft Program Feedback Report (PFR) within a certain time period. If there are no objections, the program accepts the findings written in the PFR. Once the objection period has passed, the evaluation team prepares final PFR and sends it to accreditation body.

During the site visit, facilities are examined, classes are observed, and interviews are conducted with faculty members, students, dean/principal and other relevant people (Aktan & Gencel, 2010). In addition, interviews can be conducted with university senior management, program managers, administrative staff (Özçiçek & Karaca, 2019), alumni and other stakeholders.

It is difficult to find a significant difference among the standards of the program accreditation bodies. Therefore, it can be stated that the standards of each accreditation body are similar to each other. Although there are minor differences between accreditation bodies, these standards generally consist of the following headings: students, program teaching objectives, program learning outcomes, teaching plan, teaching staff, management structure, infrastructure, institutional support, and financial resources and continuous improvement (TURAK, 2024; Aktan & Gencel, 2010, Kılıçaslan, 2020).

Accreditation standards are evaluated based on the PSER, evidence, on-site observation and interviews. In this process, the team is expected to conduct an evaluation based on objective criteria. Evaluators are expected to act impartially and make similar evaluations in similar situations. The team leader is tasked with ensuring consistency within the team and consistency among the programs when there is more than one program evaluation study. The main reason for this is that consistency in evaluations can affect competency decisions made by the accreditation agent. In addition to intra-team consistency, maintaining inter-team and inter-year consistency is not a responsibility of the accreditation body making the competency decision, but also a challenge it faces.

In the study, an analysis and investigation were carried out on the reports submitted by the accreditation evaluation teams to TURAK, which operates in the field of higher tourism education in Türkiye. Thus, beyond ensuring the relatively straightforward intra-team consistency, the study aims to draw attention to the examination of inter-team and inter-year consistency, which remains a serious challenge of many accreditation bodies. More importantly, the study also seeks to provide a model for accreditation agencies by analyzing a sample. The lack of prior research on analyzing inter-team and inter-year consistency in the literature is a key motivation to conduct this study.

Reducing the inconsistencies that may arise for various reasons in accreditation evaluations or conversely, ensuring consistency emerges as both a challenge and a responsibility. This article outlines TURAK's studies and processes related to consistency. Thus, it is also aimed to contribute to developing an effective model for consistency studies by sharing these practices with other accreditation bodies and the public.

Literature Review

Ensuring consistency in evaluations plays a critical role in helping accreditation bodies make a competency decision that accurately reflects the true status of the programs. Consistency in accreditation studies generally needs to be achieved at three levels: (1) within the team, (2) inter-team and (3) inter-year (TURAK Consistency Directive, Art.5).

Ensuring consistency within the team is primarily seen as the responsibility of the team leader and is established as a standard practice (TURAK Consistency Directive, Art. 6). On the other hand, the consistency control committee, along with the team leaders, takes part in ensuring consistency among the evaluations of different higher education programs and different years in a period, in other words, inter-teams and inter-years (TURAK Consistency Directive, Art. 7 & Art. 8).

Although it is relatively easy to ensure consistency within teams in program accreditation evaluations, it can be difficult to achieve consistency across inter-teams and inter-years. The reasons for this difficulty may be that evaluators have different perceptions of perfection, personal biases or stereotypes. Additionally, each evaluator may interpret the criteria differently. In this regard, Hash (2023) states that differences in perception between evaluators in the application of standards to programs are significant subject of criticism. On the other hand; many different reasons such as having more than one evaluator task, limited time between two evaluations, incomplete, superficial or detailed PSER, and time pressure etc. can lead to inconsistent or incompatible evaluations (Greenfield et al., 2009).Although rubrics are primarily the assessment tools used to provide feedback to the students to what degree they achieved the necessary learning outcomes (Bai et al., 2013; Gallardo, 2020; Alves et. Al, 2020), they are also recommended to utilize them for accreditation. A rubric is a qualitative instrument and criterion-based tool (Cicek et al., 2022) and

is a kind of tool used in assessment to evaluate and score a variety of tasks, assignments or projects based on predefined criteria (Jaiswal, 2024).

There are many kinds of rubrics such as analytic, holistic, single-point, multi-level, checklist, numeric, and behavioral rubrics. In TURAK assessment numeric rubrics were used. A numeric rubric requires assigning the numerical scores to performance levels; therefore, it provides a quantitative measure of achievement (Jaiswal, 2024). Rubrics may be helpful in curriculum development and evaluation, accreditation, tenure and promotion, and teacher self-evaluation (Jonsson & Svingby, 2007; Dahal, 2022; Kiesler & Impagliazzo, 2022; Kiesler & Impagliazzo, 2023; Fitriyani et al., 2024).

Research suggests that using the rubric system in evaluations can help evaluators choose the right category and also contribute to consistency when there is more than one evaluator (Shryock & Reed, 2009). On the other hand, inter-rater reliability could be used to assess compliance (Sayre-Stanhope, 2005). Since rubrics could provide some statistics, it would be beneficial to use them in accreditation evaluation process (Bishop et al., 2012).

The reliability of accreditation bodies is closely related to the compatibility or consistency of the decisions they make, which depends on the consistency maintained in the accreditation processes (Hash, 2023). In health-related accreditation studies, evaluators encounter challenges in achieving consistency in evaluation (Greenfield et al., 2009). Although consistency is very critical in accreditation evaluations, it has been found that there are very few studies addressing this issue (Hash, 2023).

Method

Information about TURAK Evaluations

TURAK bases its program accreditation evaluations on nine dimensions: 1: Students, 2: Program Teaching Objectives, 3: Program Learning Outcomes, 4: Teaching Plan, 5: Teaching Staff, 6: Management Structure, 7: Infrastructure, 8: Institutional Support and Financial Resources, and 9: Continuous Improvement. A team consisting of one team leader, one academic evaluator, one sector evaluator and one student evaluator is established for the applying programs. Prior to 2021, item-based evaluations were categorized as "Deficiency, Weakness, Anxiety and Observation". In 2021, a numeric rubric system was formed and categories were as follows: Deficit (1), Inadequate (2), Acceptable (3), Good (4), Very good (5).The meaning of each category in the numeric rubric used is as follows:

1.Deficit: Indicates that a criterion is not entirely met. The program is not in compliance with the criterion and the institution must take urgent measures to meet this criterion.

2.Inadequate: Only a small portion of a criterion is met and the institution must take measures to meet and correct this criterion.

3.Acceptable: Indicates that a criterion is partially met. The program is not in full compliance with the criterion. Measures must be taken to correct deficiencies and to meet the criterion more strongly.

4.Good: It is a statement that a criterion has been met and must be maintained.

5.Very Good: Indicates that a criterion is met very well and that some of the practices are and/or can be followed by other institutions.

The average score for a dimension is obtained by summing the average scores of all items within that dimension and dividing by the number of items. The fractions in the obtained average are converted into the corresponding numerical and qualitative value in the numeric rubric system by rounding down 0.49 and down, and rounding up 0.5 and above. Competency decisions are made by taking these numerical values into consideration.

There are four types of competency decisions, based on numerical values, made by TURAK. A brief summary of these decisions as follows (TURAK Higher Education Tourism Programs Evaluation and Accreditation Application Principles Directive, 2023, Art. 13/4/a-b-c-d):

(a) Full Accreditation: A program is granted 3+3 years of full accreditation if it meets all of the minimum conditions specified in the TURAK standards. This type of accreditation means that there is no "1: Deficient" and "2: Inadequate" evaluation in any program-related standards.

(b) Conditional Accreditation: If a program has not been evaluated as "1: Deficient" in any criterion in its overall evaluation, but has been evaluated as at least "2: Insufficient" in one or more criteria, besides 3, 4 and 5, accreditation is granted for two (2) years only. A program cannot be nominated for conditional accreditation for more than two consecutive terms.

(c) Accreditation Candidacy: Although a "1: Deficient" evaluation has not been made in any criterion in the general evaluation of a program, if "(4)" and "(5)" scoring has not been made in any of the criteria, in another words all dimensions have (2) or (3) points, and it is anticipated that the deficiencies can be completed within one year, a "Accreditation Candidate" status is granted for (1) year. A program cannot be nominated for accreditation for more than two consecutive terms.

(d) Re-Application: The decision to re-apply for accreditation means that the minimum conditions specified in the TURAK criteria are not met in the evaluation of a program evaluated for the first time. In other words, it means that a "1: Missing" evaluation is made in any criterion.

Data Set

An application was submitted to TURAK on 19.06.2023 to use the reports sent by the accreditation teams in 2021, 2022 and 2023. TURAK granted permission to use the reports with its decision dated on 22.06.2023 and numbered 5/1.A data set was created using the numerical values in the evaluation reports submitted by the teams to TURAK in 2021, 2022, 2023. A general evaluation was made in 11 programs in 2021, 9 in 2022 and 20 programs in 2023.

The number of teams formed over the years remained below 30. Therefore, based on the Central Limit Theorem, the Mann Whitney-U test was used to assess the consistency between associate and undergraduate programs. On the other hand, the Kruskal Wallis-H test was used to evaluate the consistency inter-years. The fact that there was no statistically significant difference between the averages obtained from the rubric system in the tests was interpreted as an indicator of consistency in the evaluations.

Findings

Information on Programs Applying for Accreditation

Table 1 shows the distribution of general evaluations made between 2021-2023 by university type, program level, and program type. Accordingly, there is a more or less equal distribution according to the distinction between state and foundation universities. On the other hand, it turns out that most applications are from undergraduate programs.

The distribution of general accreditation and evaluation studies conducted among the years of 2021-2023 by faculties is shown in Table 2. Accordingly, 14 of the total 40 applications are from tourism faculties. The number of undergraduate programs in total applications is 23 and the number of associate degree programs is 17.

Table 1
General Evaluation Applications by University, Program Level and Type

Criteria	Category	Year			Total
		2021	2022	2023	
University Type	State	5	4	12	21
	Foundation	6	5	8	19
	Total	11	9	20	40
Level	Bachelor	6	5	12	23
	Associate	5	4	8	17
	Total	11	9	20	40

Table 2
General Evaluation Applications by Years and Higher Education Units

Higher Education Units	Year			Total
	2021	2022	2023	
Tourism Faculty	4	3	7	14
School of Tourism & Hotel Management	0	0	1	1
Faculty of Arts & Design	0	1	1	2
School of Applied Sciences	1	1	1	3
Faculty of Business	0	0	1	1
Faculty of Arts	1	0	0	1
Faculty of Applied Sciences	0	0	1	1
Vocational School (Associate degree)	5	4	8	17
Total	11	9	20	40

Table 3 shows the distribution of the programs evaluated for the first time by years. Accordingly, among the programs receiving the most applications, the "Gastronomy and Culinary Arts" undergraduate program leads with 11 applications, followed by the associate degree "Cookery" program with 9 applications. A similar trend is observed in tourist guiding programs. The number of applications regarding tourist guiding from associate and undergraduate programs is 9. This may reflect the popularity of these programs and their tendency to both register their educational quality and use accreditation that may have at the end of the process as a promotional tool.

Table 3
General Evaluation Applications by Years and Departments

Programs	Year			Total
	2021	2022	2023	
Tourism Management	1	0	5	6
Gastronomy and Culinary Arts	3	3	5	11
Tourism Guidance	2	2	2	6
Cookery	3	2	4	9
Tourism and Hotel Management	2	1	1	4
Tourist Guidance	0	0	3	3
Tourism and Travel Services	0	1	0	1
Total	11	9	20	40

Evaluation According to Dimensions

Table 4 presents descriptive statistics for TURAK accreditation dimensions. As seen in the table, the mean scores of all dimensions, except Continuous Improvement, is above 3.00. The lowest average is observed in the continuous improvement dimension ($\bar{x}$: 2.65±0.89). On the other hand, the highest average is in Institutional Support and Financial Resources ($\bar{x}$:3.73±0.78). Therefore, programs

generally need to enhance their continuous improvement activities. Additionally, the standard deviations in terms of dimensions vary between 0.662 and 0.893. Considering that the standard deviation is close to each other and small from a 5-point numeric rubric system, it can be inferred that the consistency between teams is ensured at a minimum level.

Table 4
Descriptive Statistics According to Accreditation Criteria (n:40)

Dimensions	$\bar{x}$	Standard Deviation	Skewness	Kurtosis
1.Students	3.10	.672	.417	.730
2.Program Teaching Objectives	3.18	.747	.088	-.366
3.Program Learning Outcomes	3.27	.877	.378	-.370
4.Teaching Plan	3.40	.841	-.076	-.548
5.Teaching Staff	3.68	.797	-.292	-.152
6.Management Structure	3.60	.709	-.145	-.038
7.Infrastructure	3.65	.662	-.588	.484
8.Institutional Support and Financial Resources	3.73	.784	-.470	.131
9.Continuous Improvement	2.65	.893	.999	.774
Evaluation categories:1:Deficit, 2:Inadequate, 3:Acceptable, 4:Good, 5:Very good				

Table 5 presents some descriptive statistics for the evaluation criteria of tourism-related programs from 21 public and 19 foundation universities assessed for accreditation. Accordingly, no significant difference is obtained according to university type in any of the dimensions, except "Continuous Improvement". On the other hand, a significant difference is found in the "Continuous Improvement" dimension. It should be emphasized that this difference is in favor of foundation universities. This indicates that the consistency is maintained between teams attending public or foundation universities, with the exception for one dimension- continuous improvement.

A comparison of the programs evaluated for accreditation between 2021 and 2023 according to their types is shown in Table 6. Accordingly, there is no difference in evaluations between undergraduate and associate degree programs in all dimensions except "Continuous Improvement". This indicates that a certain level of consistency is achieved between teams depending on the level of the program. It suggests that associate degree programs perform better than undergraduate programs in terms of continuous improvement.

Table 5
Comparison of Dimensions by University Type

Dimensions	University Type	N	$\bar{x}$	Standard deviation	Z	p
1.Students	State	21	2.90	0.539	-1.898	0.105
	Foundation	19	3.32	0.749		
2. Program Teaching Objectives	State	21	2.95	0.590	-2.035	0.061
	Foundation	19	3.42	0.838		
3. Program Learning Outcomes	State	21	3.19	0.873	-0.711	0.520
	Foundation	19	3.37	0.895		
4.Teaching Plan	State	21	3.14	0.727	-2.021	0.057
	Foundation	19	3.68	0.885		
5.Teaching Staff	State	21	3.81	0.750	-1.148	0.294
	Foundation	19	3.53	0.841		
6.Management Structure	State	21	3.62	0.669	-0.299	0.789
	Foundation	19	3.58	0.769		
7.Infrastructure	State	21	3.57	0.598	-0.778	0.503
	Foundation	19	3.74	0.733		
8.Institutional Support and Financial Resource	State	21	3.76	0.700	-0.255	0.830
	Foundation	19	3.68	0.885		
9.Continuous Improvement	State	21	2.38	0.740	-2.320	0.034
	Foundation	19	2.95	0.970		
Evaluation categories:1:Deficit, 2:Inadequate, 3:Acceptable, 4:Good, 5:Very good						

Table 6
Comparison of Dimensions by Program Level

Dimensions	Level	N	$\bar{x}$	Standard Deviation	Z	p
1.Students	Bachelor	23	3.04	0.638	-0.444	0.705
	Associate	17	3.18	0.728		
2. Program Teaching Objectives	Bachelor	23	3.00	0.798	-1.892	0.085
	Associate	17	3.41	0.618		
3. Program Learning Outcomes	Bachelor	23	3.30	0.974	-0.176	0.871
	Associate	17	3.24	0.752		
4.Teaching Plan	Bachelor	23	3.39	0.839	-0.117	0.914
	Associate	17	3.41	0.870		
5.Teaching Staff	Bachelor	23	3.61	0.891	-0.416	0.705
	Associate	17	3.76	0.664		
6.Management Structure	Bachelor	23	3.61	0.783	-0.287	0.808
	Associate	17	3.59	0.618		
7.Infrastructure	Bachelor	23	3.65	0.647	0.000	1.000
	Associate	17	3.65	0.702		
8.Institutional Support and Financial Resource	Bachelor	23	3.74	0.689	-0.015	1.000
	Associate	17	3.71	0.920		
9.Continuous Improvement	Bachelor	23	2.35	0.714	-2.537	0.019
	Associate	17	3.06	0.966		
Evaluation categories:1:Deficit, 2:Inadequate, 3:Acceptable, 4:Good, 5:Very good						

A "Re-Application" decision was made for one program from 2021 to 2023. Except for this program, "Conditional Accreditation" and "Full Accreditation" decisions have been made for other programs. The comparison made according to the type of accreditation received as a result of the general evaluation is presented in Table 7. It is expected that statistically significant differences exist, with the average scores of programs receiving "Full Accreditation" being higher than those granted "Conditional Accreditation.

Table 7 presents that there is not any significant difference in the dimensions of "Teaching Staff", "Infrastructure" and "Institutional Support and Financial Resources". There are similar evaluations in these dimensions. On the other hand, there are significant differences in the dimensions of "Students", "Program Teaching Objectives", "Program Learning Outcomes", "Teaching Plan", "Management Structure" and "Continuous Improvement". It is observed that in these dimensions, the average scores in programs receiving full accreditation are higher than in programs receiving conditional accreditation.

The dimensions for which no statistical difference was obtained generally belong to issues beyond the control of program managers. On the other hand, the dimensions in which significant differences were found are predominantly the dimensions in which the program can control and influence.

Excluding one program for which a "Re-Application" decision was made, the relationships between the decisions made by TURAK, the university type and program level were examined by creating a cross-tabulation. According to Table 8, there is no relationship between the level of the program and the accreditation decision.

Table 9 shows a relationship between university type and accreditation decision. Accordingly, it is determined that there is a significant relationship between the type of university and the accreditation decision. While 18 of 21 state universities (85.7%) were conditionally accredited, 9 of 18 foundation universities (50%) received full accreditation. From another perspective, 18 of the 27 conditionally accredited programs (66.7%) are in state universities. On the other hand, 9 of the 12 programs that received full accreditation (75%) are in foundation universities. The Phi coefficient is 0.386 indicates a low-level relationship between the type of university and the accreditation received (Alpar, 2010, p. 277).

Table 7
Comparison of Dimensions Based on the Accreditation Decisions

Dimensions	Accreditation Assigned	N	$\bar{x}$	Standard Deviation	Z	p
1.Students	Conditional accreditation	27	2.78	0.424	-4.570	p<0.0001
	Full accreditation	12	3.83	0.577		
2. Program Teaching Objectives	Conditional accreditation	27	2.81	0.557	-4.654	p<0.0001
	Full accreditation	12	4.00	0.426		
3. Program Learning Outcomes	Conditional accreditation	27	2.93	0.616	-4.094	p<0.0001
	Full accreditation	12	4.17	0.718		
4. Teaching Plan	Conditional accreditation	27	3.15	0.718	-3.242	0.002
	Full accreditation	12	4.08	0.669		
5. Teaching Staff	Conditional accreditation	27	3.59	0.747	-1.449	0.188
	Full accreditation	12	4.00	0.739		
6. Management Structure	Conditional accreditation	27	3.41	0.572	-3.215	0.003
	Full accreditation	12	4.17	0.577		
7. Infrastructure	Conditional accreditation	27	3.59	0.572	-1.346	0.258
	Full accreditation	12	3.92	0.669		
8. Institutional Support & Financial Resource	Conditional accreditation	27	3.59	0.797	-1.745	0.118
	Full accreditation	12	4.08	0.669		
9. Continuous Improvement	Conditional accreditation	27	2.33	0.679	-4.332	p<0.0001
	Full accreditation	12	3.50	0.674		
Evaluation categories:1:Deficit, 2:Inadequate, 3:Acceptable, 4:Good, 5:Very good						

Table 8
Relationship Between Program Level and Accreditation Decisions

Level	Statistics	Accreditation Assigned		Total
		Conditional	Full	
Associate	f	16	6	22
	% in level	72.7	27.3	100.0
	% in accreditation assigned	59.3	50.0	56.4
Bachelor	f	11	6	17
	% in level	64.7	35.3	100.0
	% in accreditation assigned	40.7	50.0	43.6
Total	f	27	12	39
	% in level	69.2	30.8	100.0
	% in accreditation assigned	100.0	100.0	100.0
	f	16	6	22
Fisher Exact test: p (2-sided): 0.730; Minimum expected value: 5.23; ϕ (Phi):0.086; p:0.590				

Table 9
Relationship between University Type and Accreditation Decisions

University Type	Statistics	Accreditation Assigned		Total
		Conditional	Full	
State	f	18	3	21
	% in university	85.7	14.3	100.0
	% in accreditation assigned	66.7	25.0	53.8
Foundation	f	9	9	18
	% in university	50.0	50.0	100.0
	% in accreditation assigned	33.3	75.0	46.2
Total	f	27	12	39
	% in university	69.2	30.8	100.0
	% in accreditation assigned	100.0	100.0	100.0
	% in total	69.2	30.8	100.0
Fisher Exact test: p (2-sided): 0.035; Minimum expected value: 5.54; ϕ (Phi):0.386; p:0.016				

Table 10
Comparison of Dimensions by Years

Dimensions	Year	N	$\bar{x}$	Standard deviation	Kruskal-Wallis H (df; p value)
1.Students	2021	11	3.18	.874	0.092; (2; 0.955)
	2022	9	3.11	.333	
	2023	20	3.05	.686	
2.Program Teaching Objectives	2021	11	3.45	.688	2.732 (2; 0.255)
	2022	9	3.11	.601	
	2023	20	3.05	.826	
3.Program Learning Outcomes	2021	11	3.27	1.009	4.153 (2; 0.125)
	2022	9	2.78	.667	
	2023	20	3.50	.827	
4.Teaching Plan	2021	11	3.45	1.036	0.661 (2; 0.719)
	2022	9	3.22	.667	
	2023	20	3.45	.826	
5.Teaching Staff	2021	11	3.73	.786	0.042 (2; 0.979)
	2022	9	3.67	.707	
	2023	20	3.65	.875	
6.Management Structure	2021	11	3.64	.809	3.332 (2; 0.189)
	2022	9	3.22	.667	
	2023	20	3.75	.639	
7.Infrastructure	2021	11	3.91	.701	4.605 (2; 0.100)
	2022	9	3.33	.707	
	2023	20	3.65	.587	
8.Institutional Support & Financial Resource	2021	11	4.09	.701	3.297 (2; 0.192)
	2022	9	3.56	.527	
	2023	20	3.60	.883	
9.Continuous Improvement	2021	11	3.09	1.221	1.777 (2; 0.411)
	2022	9	2.44	.882	
	2023	20	2.50	.607	
Evaluation categories:1:Deficit, 2:Inadequate, 3:Acceptable, 4:Good, 5:Very good					

Evaluation by Years

The Kruskal Wallis H test was used to determine whether there was a significant difference by year for the 40 programs that were evaluated for the first time from 2021 to 2023 (Table 10). Accordingly, there is no statistically significant difference in any of the 9 dimensions in TURAK evaluations. This indicates that consistency has been achieved inter- years in terms of general evaluations.

Discussion

Ensuring consistency inter-teams and inter-years is a significant challenge for accreditation bodies. Inconsistencies in this regard may be caused by evaluators' perception differences (Hash, 2023), while limited time pressure may also lead to differences in evaluations (Greenfiel et al., 2009).

Given the fact that consistency is an important requirement (Tierney & Simon, 2004),the rubric system may provide an advantage in ensuring consistency in evaluations (Shryock & Reed, 2009). TURAK teams make evaluations according to the rubric system. Based on the claims of Jonsson and Svingby (2007), we can assume that the more consistent the scores over different teams, years, universities and program levels, the more reliable the assessment.

The use of rubrics in evaluations can pave the way for statistical analysis due to the possibility of quantification. This allows consistency to be measured statistically. In this direction, the use of holistic and summative rubrics (Wilkerson, 2019) becomes important.

Essentially, the use of rubrics, which are effective tools for the evaluation of learning outcomes, can provide significant convenience for evaluators, consistency control committees and accreditation bodies in program accreditation. The most important benefit for evaluators is the ease of reduction from an abstract level to quantification. However, different perceptions of excellence, misunderstandings or evaluators' workloads can lead to inconsistencies in the evaluation of accreditation criteria according to the specified rubric. In order to prevent this, a specified rubric can be used for each criterion, which can lead to complexity and a decrease in the demand for voluntary evaluation.

The findings imply that foundation universities are better than state universities in terms of continuous improvement. This can be explained by the fact that the management style in foundation universities is more flexible than state universities. Also, the centralized structure in state universities can be another reason.

The findings imply that foundation universities are better than state universities in terms of continuous improvement. This can be explained by the fact that the management style in foundation universities is more flexible than state universities. Also, the centralized structure in state universities can be another reason (Kurtay & Duran, 2018). On the other hand, a significant difference is detected when the significance level is taken as 10% in the program learning outcomes. This finding may be explained as foundation universities are more commercial or profit-oriented. The fact that foundation universities are characterized as demand absorbers (Kwiek, 2011), in other words, that they develop their programs according to demand, may have paved the way for this difference.

Research findings reveals that assessment by using rubrics has shown statistically significant difference based on the accreditation granted, i.e. full or conditional accreditation. These differences had been found out especially on students' enrollment and services, program teaching objectives, program learning outputs, teaching plan, management structure, and continuous improvement. As Cura and Alani explained (2018) , we could claim that Once the accreditation process begun and became an internalized organizational culture, improvements in programs were inevitable.

Conclusion

The evaluations of the programs applying to TURAK for accreditation for the first time from 2021 to 2023 were compared based on the university types, the level of the program, the type of the program and the years. TURAK evaluations are carried out on 9 dimensions. The standard deviations of the nine dimensions, calculated on a 5-point scale, are relatively small. These deviations range from 0.672 to 0.893.This indicated that a minimum level of consistency among the teams was achieved.

When comparing evaluations across the nine dimensions between state and foundation universities, any significant difference was not obtained in eight dimensions, except for 'Continuous Improvement.' This indicates that inter-team consistency for state and foundation universities was predominantly managed. Similarly, when comparing evaluations based on the program level, any significant difference was not found, except for "Continuous Improvement". This indicates that there is not any notable problem in terms of inter-team consistency between the associate and undergraduate programs.

In a comparison based on the accreditation decisions made by TURAK, "Conditional Accreditation" and "Full Accreditation", after the team reports and the consistency control committee's review, it was revealed that there was a significant difference in 6 out of the 9 dimensions. Such a significant difference is expected in this type of comparison. On the other hand, the fact that no significant difference was found in the dimensions- "Teaching Staff", "Infrastructure" and "Institutional Support and Financial Resources"- implies a situation caused by factors beyond the control of program managers. However, it is determined that there are significant differences in the dimensions-

"Students", "Program Teaching Objectives", "Program Learning Outcomes", "Teaching Plan", "Management Structure" and "Continuous Improvement"- which can be manageable and reflect the impact of the departmental management effectiveness. The averages in programs receiving full accreditation in these dimensions are higher compared to programs receiving conditional accreditation. This fact suggests those department managers, academic and administrative staffs working in the department have positive contributions to the accreditation process.

Ensuring inter-year consistency in accreditation studies is one of the most vital concerns. The analysis revealed that the evaluations made by TURAK teams in terms of 9 dimensions were not significantly different. This situation reveals that inter-year consistency, which is a critically important concern, has been achieved at a certain level.

In TURAK evaluations, the evaluation differences for the items in the criteria should not exceed 2.00, which make an important contribution to ensuring consistency within the team. The findings confirm that, as stated by Shryock and Reed (2009), and highlighted by Shabani and Panahi (2020)the use of a rubric system in accreditation evaluations is an effective tool in ensuring inter-team and inter-year consistency. Additionally, as highlighted by Shabani and Panahi (2020) rubrics could help evaluators to be consistent in assessment and improve validity. Therefore, it can be suggested that accreditation bodies should use the rubric system in evaluations. However, it is a fact that an analytical evaluation based on solely learning outcomes is not sufficient for accreditation. Therefore, it is necessary to address the evaluation dimensions in accreditation in a holistic manner, including rubric-based evaluation.

As mentioned before, rubrics improve student learning outcomes, teaching effectiveness, and course design enhance self-assessment (Ragupathi & Lee, 2020). Therefore, the use of rubrics in accreditation processes can contribute the most to the evaluation of learning outcomes. On the other hand, developing rubrics for other criteria remains an area with potential for further exploration. Therefore, in subsequent studies, it would be beneficial to focus on developing rubrics for criteria other than learning outcomes.

References

Aktan, C. C., & Gencel, U. (2010). Accreditation in higher education. *Journal of Organization and Management Sciences*, *2*(2), 137-146.

Alpar, R. (2010). *Uygulamalı istatistik ve geçerlik-güvenirlik*. Detay Yayıncılık.

Alves, N. D. C., von Wangenheim, C. G., Hauck, J. C. R., & Borgatto, A. F. (2020). *A large-scale evaluation of a rubric for the automatic assessment of algorithms and programming concepts*. In Proceedings of the 51st ACM Technical Symposium on Computer Science Education (pp. 556-562). https://doi.org/10.1145/3328778.3366840

Arnanz, J., & Kaewnuch, K. (2019). Accreditation in tourism and hospitality undergraduate education in the ASEAN context: The case of Thailand from the Ted Qual perspective. *The Journal of Social Sciences Research*, *5*(2), 275-283. https://doi.org/10.32861/jssr.52.275.283

Bai, X., Xu, Y., & Ikem, F. (2013). Rubric and performance-based assessment. *Issues in Information Systems*, *14*(2). https://doi.org/10.48009/2_iis_2013_1-11

Bishop, W., Nespoli, O., & Parker, W. (2012). *Rubrics for accreditation and outcomes assessment in engineering capstone projects*. In Proceedings of the Canadian Engineering Education Association (CEEA).

Çiçek, J. S., Mann, D., & Renaud, R. (2022). Supporting teaching practice, program improvement, and accreditation efforts in an engineering program. *Canadian Biosystem Engineering Journal*, *64*, 9.1-9.11. https://doi.org/10.7451/CBE.2022.64.9.1

Cura, F., & Ahmed Alani, T. (2018). Accreditation effect on quality of education at business schools. *International Journal of Social Sciences & Educational Studies*, *4*(5), 71-82. https://doi.org/10.23918/ijsses.v4i5p71

Dahal, S. (2022). *Using rubrics in continuous assessment system: A narrative inquiry* [Yayımlanmamış doktora tezi]. Kathmandu University.

Fitriyani, N., Evendi, E., & Suwanto, S. (2024, Mart). *The effect of using rubrics in improving the quality of assessment of mathematics learning*. In Proceedings of International Seminar on Student Research in Education, Science, and Technology, 1(April 2024), 91-101.

Gallardo, K. (2020). Competency-based assessment and the use of performance-based evaluation rubrics in higher education: Challenges towards the next decade. *Problems of Education in the 21st Century*, *78*(1), 61-79. https://doi.org/10.33225/pec/20.78.61

Greenfield, D., Pawsey, M., Naylor, J., & Braithwaite, J. (2013). Researching the reliability of accreditation survey teams: Lessons learned when things went wrong. *Health Information Management Journal*, *42*(1), 4-10.

Hash, R. (2019). Consistency in decision-making between survey teams and the decision-making body in a professional education program accrediting agency [Versiyon 1]. *Med Ed Publish*, *8*, 113. https://doi.org/10.15694/mep.2019.000113.1

Jaiswal, C. P. (2024). Modern method of curriculum. Assessment: Rubrics. *Anusandhan Anveshika*, *XIV*(Temmuz 2024), 21-30.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130-144. https://doi.org/10.1016/j.edurev.2007.05.002

Kılıçaslan, Ç. (2020). Higher education accreditation. PEYZAJ - *Journal of Education, Science, Culture and Art Special Issue*, 10-18.

Kiesler, N., & Impagliazzo, J. (2022). *Implementing diversity, equity, and inclusion in accreditation*. In Proceedings of International Symposium on Accreditation of Engineering and Computing Education (ICACIT) (pp. 1-6). IEEE. https://doi.org/10.1109/ICACIT56139.2022.10041468

Kiesler, N., & Impagliazzo, J. (2023). *Evaluating diversity, equity, inclusion, and accessibility in accreditation*. In Proceedings of International Symposium on Accreditation of Engineering and Computing Education (ICACIT) (pp. 1-5). IEEE. https://doi.org/10.1109/ICACIT59946.2023.10403684

Kumar, P., Shukla, B., & Passey, D. (2020). Impact of accreditation on quality and excellence of higher education institutions. *Investigación Operacional*, *41*(2), 151-167.

Kurtay, M. Z., & Duran, A. (2018). Vakıf üniversitesinde akademisyen olmak: Rutinler ve bu rutinleri etkileyen faktörler. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *18*(4), 2518-2540.

Kwiek, M. (2011, March). *The public/private dynamics in Polish higher education: Demand-absorbing private growth and its implications*. In Proceedings of Higher Education Forum, 8 (March), 37-59.

Özçiçek, Y., & Karaca, A. (2019). Quality and accreditation in higher education institutions: Evaluation of engineering education programs. *Fırat University International Journal of Economics and Administrative Sciences*, *3*(1), 114-149.

Prøitz, T. S., Stensaker, B., & Harvey, L. (2004). Accreditation, standards and diversity: An analysis of EQUIS accreditation reports. *Assessment & Evaluation in Higher Education*, *29*(6), 735-750. https://doi.org/10.1080/0260293042000227263

Ragupathi, K., & Lee, A. (2020). Beyond fairness and consistency in grading: The role of rubrics in higher education. In C. S. Sanger & N. W. Gleason (Eds.), *Diversity and inclusion in global higher education: Lessons from across Asia* (pp. 73-95). Palgrave Macmillan. https://doi.org/10.1007/978-981-15-1628-3

Sayre-Stanhope, D. (2005). Consistency of review of accreditation documents by commissioners of the Accreditation Review Commission on Education for the Physician Assistant. *The Journal of Physician Assistant Education*, *16*(2), 73-78.

Semerci, C. (2017). Development of the perception of accreditation (PA) scale: Validity and reliability study. *Bartın University Faculty of Education Journal*, *6*(3), 1093-1104.

Shabani, E. A., & Panahi, J. (2020). Examining consistency among different rubrics for assessing writing. *Language Testing in Asia*, *10*, 1-25. https://doi.org/10.1186/s40468-020-00111-4

Shryock, K., & Reed, H. (2009). ABET *accreditation-best practices for assessment. American Society for Engineering Education*, In Proceedings of 2008 ASEE Gulf-Southwest Annual Conference. The University of New Mexico – Albuquerque.

Sin, C., Tavares, O., & Amaral, A. (2017). The impact of programme accreditation on Portuguese higher education provision. *Assessment & Evaluation in Higher Education*, *42*(6), 860-871. https://doi.org/10.1080/02602938.2016.1203860

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research, and Evaluation*, *9*(1). https://doi.org/10.7275/jtvt-wg68

TURAK. (2020). *Consistency guideline*. https://drive.google.com/file/d/1aZAde2PbMX2-fQOLUnJ3gWikgXq-v6Vy/view

TURAK. (2023). *Higher education tourism programs evaluation and accreditation application principles directive*. https://docs.google.com/document/d/16asbTD6gHezNuXgn1gEXaiYiCRRgB9If/edit

Wilkerson, J. R. (2020). Rubrics meeting quality assurance and improvement needs in the accreditation context. *Quality Assurance in Education*, *28*(1), 19-32. https://doi.org/10.1108/QAE-04-2019-0045

Yükseköğretim Kalite Kurulu. (2018, Kasım 23). Yükseköğretim Kalite Güvencesi ve Yükseköğretim Kalite Kurulu Yönetmeliği. *T.C. Resmî Gazete*, Sayı 30604. https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=28996&MevzuatTur=7&MevzuatTertip=5