International Journal of Assessment Tools in Education
2026, Vol. 13, No. 1, 1–20
https://doi.org/10.21449/ijate.1610075
journal homepage: https://dergipark.org.tr/en/pub/ijate
Research Article

Assessing feedback practices in higher education: Development and validation of the Perceived Quality of Feedback Scale

Emrullah Esen[1*], Oktay Cem Adıgüzel[1], Derya Atik Kara[1]

[1]Anadolu University, Faculty of Education, Department of Educational Sciences, Eskişehir, Türkiye

ARTICLE HISTORY

Received: Dec. 30, 2024

Accepted: May 29, 2025

Keywords:

Feedback quality,
Functions of feedback,
Content of feedback,
Presentation of feedback,
Scale development.

Abstract: Despite the importance of feedback quality in educational contexts, the availability of comprehensive, valid, and reliable tools for assessing feedback quality remains limited. The objective of this study was to develop a valid and reliable measurement tool for evaluating the perceived quality of feedback in a comprehensive manner. The study was conducted with the participation of 847 pre-service teachers from a range of grade levels and teacher education programmes (322 second-year, 272 third-year, 253 fourth-year; 52.18% female, 47.82% male) studying at a public university's Faculty of Education. Participants were selected from various teaching programs. The scale was developed in three stages: item development, scale development, and scale evaluation. Exploratory and confirmatory factor analyses were performed to examine the factor structure of the assessment tool. Exploratory factor analysis (EFA) was conducted with data collected from 531 undergraduate students, followed by confirmatory factor analysis (CFA) with data from another group of 316 undergraduate students. The EFA results revealed a five-factor structure consistent with the theoretical framework: cognitive orientation, metacognitive orientation, affective orientation, quality of presentation, and quality of content. These five factors explained 52.53% of the total variance. The CFA demonstrated that the model's fit indices exceeded the acceptable thresholds (CFI=.95, RMSEA=.05, TLI=.94). The reliability analyses showed that the Cronbach's Alpha (.78 - .93) and McDonald's Omega (.77 - .93) coefficients for the scale dimensions were within acceptable ranges, confirming the scale's reliability. Consequently, a valid and reliable 25-item scale with five dimensions was developed to measure the perceived quality of feedback.

INTRODUCTION

Feedback is a fundamental tool for the learning-teaching process and has long been of great interest to many researchers. The role of feedback in educational theory has been a topic of considerable interest, with numerous perspectives having been adopted in the analysis of its effects on learning. The importance of feedback for cognitive performance is well documented (e.g., Hattie & Timperley, 2007; Shute, 2008). However, studies also highlight its value for the development of metacognitive and affective characteristics (Holstun *et al.*, 2021; Panadero & Lipnevich, 2022; Prilop *et al.*, 2021). Nevertheless, despite its multifaceted significance, a substantial body of research indicates that feedback practices are marred by significant

*Corresponding author at: Emrullah ESEN emrullahesen@anadolu.edu.tr Anadolu University, Faculty of Education, Department of Educational Sciences, Eskişehir, Türkiye

e-ISSN: 2148-7456

shortcomings at the local and international level (e.g., Henderson *et al.*, 2019; Williams, 2024). The problems associated with feedback are multidimensional, involving the quality of feedback (Nugraheny *et al.*, 2016; Ostaeyen *et al.*, 2023), dissatisfaction with feedback (Deeley *et al.*, 2019), contextual barriers (Henderson *et al.*, 2019), engagement with feedback (Winstone *et al.*, 2017), and lack of empirical guidelines and strategy (O'Donovan *et al.*, 2016). Students and educators face challenges in generating and understanding effective feedback, with attitudes and skills potentially obstructing this process (Henderson *et al.*, 2019). Despite its potential to significantly influence student learning and performance, feedback often fails to achieve this potential due to various factors.

First of all, studies underline the problem of the 'feedback gap' that the actual quality of feedback falls behind the established quality standards. In a comprehensive study conducted with 149 students in the field of health sciences, Ostaeyen *et al.* (2023) revealed striking findings that only 14.9% of the feedback was of high quality, the majority was of medium quality, and the elaboration criterion was hardly met. Similarly, in their focus groups with 44 undergraduates in the social sciences, Deeley *et al.* (2019) found that students perceived feedback as inconsistent and inadequate. Glazzard and Stones (2019) revealed the main problems in feedback practices in a focus group study with seven pre-service teachers at a university in England. The study clearly demonstrated the limitations of feedback modes, the inadequacy of written feedback, and students' difficulty in comprehending feedback

Secondly, studies have shown perceptual differences in feedback practices across higher education contexts. Mulliner and Tucker's (2017) quantitative research with 26 lecturers and 194 undergraduates revealed striking differences in perception; 93% of students claimed to read feedback, while only 35% of lecturers believed this. Additionally, 85% of lecturers rated the quality of feedback positively, while only 47% of students shared the same view. Similarly, Winstone *et al.* (2017), in a focus group study with 31 psychology undergraduate students, examined in detail the psychological barriers to students' understanding, interpretation, and application of feedback and showed that students have significant difficulties receiving it. Williams's (2024) systematic review of existing research on feedback practices in higher education further illuminated significant differences between the way lecturers provide feedback and the way students receive and interpret it. The findings of these studies consistently reveal critical limitations in the provision of feedback within higher education. Feedback is often perceived as inconsistent, lacking in detailed guidance, failing to engage students effectively, and failing to translate into meaningful learning improvements. These findings point to systemic problems in the conceptualization, presentation, and reception of feedback in academic settings.

Similar challenges with feedback have been observed in studies conducted within the Turkish higher education context. Research highlights that pre-service teachers face significant obstacles in accessing feedback, with difficulties stemming from both the feedback process and its delivery (Kumral & Saracaloğlu, 2011; Öntaş & Kaya, 2019; Şahin, 2015). A qualitative case study undertaken by Kumral and Saracaloğlu (2011) at a state university revealed that the expectations for feedback among pre-service teachers were not systematically met. Qualitative analysis of focus group discussions and individual interviews with 30 pre-service teachers and four academic staff members revealed that insufficient feedback practices within the educational process resulted in feedback resistance among the pre-service teacher participants. A descriptive study by Şahin (2015) indicated notable deficiencies in feedback practices among 377 pre-service teachers in pedagogical formation certificate programs. The study indicated that verbal feedback was primarily employed, whereas written, visual, and nonverbal feedback methods were utilized to a significantly lesser degree. The study indicated that feedback was primarily teacher-centered, with minimal involvement from peers or self-assessment. Similarly, a qualitative descriptive study by Öntaş and Kaya (2019), conducted with 56 primary pre-service teachers, revealed that feedback provided in courses was not detailed enough, highlighting the strong need for specific information on their work. These studies highlight that feedback practices in teacher education programmes have significant shortcomings. These studies underline the need for systematic improvement of feedback practices in the Turkish higher education context.

The resolution of these issues can facilitate the generation of more meaningful and effective feedback, which in turn can enhance learning outcomes. While the aforementioned problems to related to feedback commonly observed across educational settings, these problems may vary according to the environment and context. This is because the feedback needs of student groups may differ and thus require different solutions. Nevertheless, in order to address these unique problems related to feedback, it is necessary that these issues should be identified in an effective, expedient, and reliable manner, and that solutions are subsequently established.

To accomplish this, it is necessary to utilize a valid, reliable, and effective measurement tool, as the inability to measure something precludes the possibility of improvement. Therefore, this study seeks to create a valid and reliable instrument for assessing the quality of feedback in a comprehensive manner, encompassing all its constituent dimensions. In doing so, the study aimed to provide a robust tool that could facilitate both educators and researchers in identifying strengths and weaknesses in feedback practices. This instrument is intended to support the development of more effective feedback processes, promote evidence-based improvements in educational environments, and ultimately enhance educational outcomes by addressing the complex nature of feedback quality.

The Issues With the Current Feedback Scales

The literature review yielded eight scales in both English and Turkish related to feedback. Two of the scales are Turkish adaptations of the original English scales. The authors conducted a thorough examination of their framework, factors, sampling methods, psychometric qualities, items, and suitability (see Table 1). ‘The Perceived Teacher Feedback Scale’, created by Koka and Hein (2003) and subsequently adapted into Turkish by Kara *et al.* (2018), was designed for primary school students and focuses on the psycho-motor characteristics of students. 'Teacher Feedback Evaluation Scale', developed by Ocak and Karafil (2020), is designed to assess the perception of feedback among secondary school students. These two scales were determined to be inappropriate for the intended population of university students.

‘Feedback Orientation Scale’, originally developed by Yang *et al.* (2014) and later adapted to Kartol and Arslan (2021), was specifically designed for the university student population. However, upon detailed examination of the factors and items, it was decided that it had limitations in comprehensively assessing the practicality of feedback regarding cognitive, affective, and metacognitive dimensions.

The scale developed by Akkuzu and Uyulgan (2014) is exclusively concerned with the dimensions of professional development and anxiety. These dimensions do not fully reflect the purpose, orientation, presentation, and quality of the content of the feedback in a holistic manner. Besides, the study includes items only relevant to the Teaching Practice Course. Similarly, the scale developed by Baydas Onlu *et al*. (2022) addresses the affective impact of feedback but does not comprehensively evaluate the cognitive, metacognitive, and affective orientations, and contains items tailored specifically to the Instructional Technologies and Material Design course.

Although the scale developed by Beydogan (2016) includes dimensions such as quality and competence in feedback, the scale does not directly focus on the function of the feedback, the quality of presentation, and the content. While Beydogan's (2016) scale effectively measures feedback characteristics, it does not address metacognitive and affective dimensions separately.

Table 1. *Current scales for measuring feedback practices.*

Title	Authors	Theoretical Framework	Factors	Sample	Psychometric Properties	Suitability for the Population	Limitations
"Perceived Teacher Feedback Scale: The Validity And Reliability Study"	Kara *et al.* (2018) Turkish adaptation of Koka & Hein, 2003	Cognitive Evaluation Theory (Deci & Ryan, 1985)	• Positive Non-verbal feedback • Positive General Feedback • Negative Non-verbal Feedback • Knowledge of Performance	454 primary and secondary school students	CFA reported in detail	Not suitable for university students	• Relevant for 10-16 age group students
"Teacher Feedback Use Evaluation Scale"	Ocak & Karafil (2020)	Based on empirical studies	• Functionality of Feedback • Structure of Feedback • Negative Feedback Structure	220 high school students	EFA and CFA reported in detail	Not suitable for university students	• Relevant for high school students • Lacks in specific theoretical framework
"Feedback Orientation Scale"	Kartol & Arslan (2021) Adaptation of Yang *et al.*, 2014	London & Smither (2002) Feedback Orientation Model	• Feedback Utility • Feedback Accountability • Feedback Social Awareness • Feedback Self-Efficacy	219 university students	Only CFA reported in detail	Suitable for university students	• Not relevant for determining cognitive, meta-cognitive and affective orientation of feedback separately due to the factor structure
"Feedback Scale"	Akkuzu & Uyulgan (2014)	Based on empirical studies	• Professional Development • Anxiety	512 pre-service from diverse programs	EFA and CFA reported in detail	Suitable for university for students	• Relevant for the Teaching Practice Course • Lacks in specific theoretical framework
"Students' Perception of Instructional Feedback Scale"	Baydas Onlu *et al.* (2022)	Based on empirical studies	• Mastery • Positive Affect • Negative Affect	231 pre-service teachers studying Computer Education and Instructional Technology	EFA and CFA reported in detail	Suitable for university for students	• Relevant for the Instructional Technologies and Material Design Course • Lacks in specific theoretical framework
"Feedback Correction Perception Scale For Teacher Candidates"	Beydogan (2016)	Based on empirical studies	• Principles in Feedback • Quality in Feedback • Style in Feedback • Sufficiency in Feedback • Scope in Feedback • Performance Orientation	670 pre-service from diverse programs	EFA and CFA reported in detail	Suitable for university for students	• Not relevant for determining cognitive, meta-cognitive and affective orientation of feedback separately due to the factor structure • Lacks in specific theoretical framework

To sum up, existing scales are insufficient for providing a comprehensive evaluation of feedback in terms of its function, presentation, and content under specific circumstances. This highlights the necessity for a scale that assesses the perceived quality of feedback from a holistic perspective within a particular theoretical framework.

Current Study

The study aimed to construct a valid, reliable, and comprehensive tool for assessing perceptions of feedback quality. In determining the structure of the scale, the effective feedback framework posited by Narciss and Huth (2004) was taken as a basis, as this framework provides a comprehensive and multidimensional framework for how feedback should be presented in terms of its function, content, and presentation. As posited by Narciss and Huth (2004), the efficacy and caliber of feedback can be attributed to three fundamental dimensions: function, presentation, and content. These dimensions are shaped by individual and instructional factors. Accordingly, the items were generated in compliance with the functional orientation, presentation, and content of feedback. Accordingly, five distinct dimensions were delineated in a deductive manner (Boateng *et al.*, 2018): ‘cognitive orientation’, ‘metacognitive orientation’, ‘affective orientation’, ‘quality of presentation’, and ‘quality of content’.

i. Cognitive orientation pertains to how feedback is provided, intending to facilitate students' cognitive processes and performance. This dimension pertains to the intention of feedback to facilitate students' cognitive processes and performance. For example, cognitive-oriented feedback might include statements such as: “*Your literature review effectively synthesizes multiple theoretical perspectives but lacks critical comparison between contradicting studies. Consider analyzing how Smith's (201X) findings challenge the conclusions drawn by Jones (201X) regarding student engagement*.” This type of feedback directly targets students' cognitive processes by highlighting analytical gaps and suggesting specific approaches to deepen their understanding of the subject matter.

ii. Metacognitive orientation pertains to the intention of feedback to facilitate students' organization of the learning process. Metacognitive-oriented feedback aims to develop students' self-regulation and awareness of their learning processes. For instance: “I noticed you completed this assignment at the last minute, which affected the depth of your analysis. For your next paper, try creating an outline two weeks before the deadline and reviewing your draft with a peer. This approach would help you identify gaps in your argument earlier in the process.” This type of feedback encourages students to reflect on their learning strategies and suggests specific techniques to monitor and improve their approach to academic tasks.

iii. Affective orientation pertains to the dimension of feedback that is intended to support students' emotional and affective processes. This dimension pertains to the intention of feedback to bolster students' motivation, self-efficacy, and attitudes. An example would be: “Your persistence in tackling this challenging statistical analysis demonstrates remarkable growth since your first assignment. The way you methodically worked through each step shows a genuine commitment to mastering these concepts. This approach will serve you well in future research projects.” Such feedback acknowledges emotional aspects of learning, reinforces positive behaviors, and builds self-efficacy even when students face difficult material.

iv. The dimension of quality of presentation pertains to the timing, frequency, mode of presentation, and style of feedback. The quality of presentation concerns how feedback is delivered. For example, a high-quality presentation might involve: “Let's review your essay draft during office hours next Tuesday at 2pm. I'll prepare specific suggestions about your argument structure, and we can discuss them face-to-face. Bring your outline and sources so we can reference them during our discussion.” This feedback specifies timing, mode of delivery, and establishes clear expectations for the feedback interaction, making it more likely that students will engage meaningfully with the feedback.

v. The dimension of quality of content pertains to the detail of feedback regarding content, information, and relevance to the student and the course needs. Quality of content refers to the substance and relevance of feedback information. For instance: “Your methodology section needs revision in three specific areas: (1) your participant selection process doesn't align with your research questions, (2) your data collection timeline is inconsistent with your stated longitudinal approach, and (3) your described analytical procedures don't adequately address potential confounding variables. Review Chapter 8 in your research methods textbook for guidance on addressing these issues.” This feedback provides detailed, specific content that directly addresses course objectives and student needs, offering clear direction for improvement."

This comprehensive framework guided authors to form a general scheme of scale and item development process. Furthermore, the theoretical framework, with its comprehensive structure, reflects the original and scientific value of the scale and the study.

METHOD

Scale Development Process

The scale development process, comprising three stages and nine steps as proposed by Boateng *et al.* (2018), was adhered to throughout the scale's development (Figure 1). At the administration of the scale, the study incorporated cross-sectional survey method in order to gather data to conduct factor analysis to validate the scale.

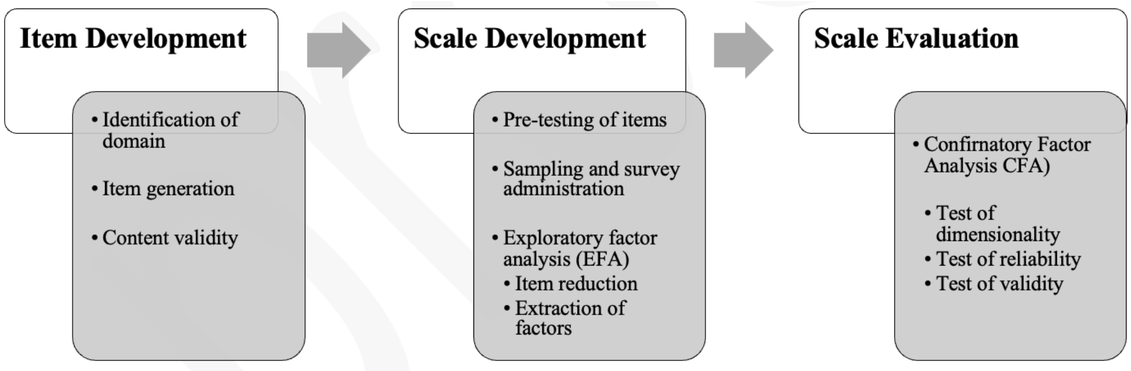

Figure 1. The guideline for the scale development (Boateng *et al.*, 2018).

At the preliminary phase of item development, the authors specified the purpose and identified the domain of the scale. Then, the authors specified dimensions of the scale a priori grounded in the theoretical framework posited by Narciss and Huth (2004) and generated 37 initial items in a “deductive way” by conducting a literature review in line with the framework utilizing the comprehensive feedback guidelines (Brookhart, 2008; Hattie & Timperley, 2007; Shute, 2008).

The 5-dimensional draft scale, comprising 37 items and 5 categories, was presented for assessment by a panel of 12 experts. The scales were distributed to seven experts in Curriculum and Instruction, three in Measurement and Evaluation, and two in Psychological Counselling and Guidance, all of whom held a doctorate in Educational Sciences. Following the expert opinions, the content validity ratio (CVR) and content validity index (CVI) values for each item were calculated (see Appendix 1). The calculations showed that the items exceed the minimum CVR ratio (.56) specified for 12 experts (Lawshe, 1975). Accordingly, the requisite amendments were implemented in the items comprising the scale, and the items were retained. The calculated CVI value for the scale is 0.95, which is above the accepted threshold (Lawshe, 1975).

At the first stage of scale development, the pilot study was conducted with the participation of 25 pre-service teachers. During the pilot study, researchers employed individual interviews to evaluate item clarity and comprehensibility while simultaneously collecting data on participants' response processes and cognitive strategies. After the pilot study, the exploratory factor analysis was implemented with the involvement of 531 pre-service educators from different teaching programmes and different grade levels. In this stage, item reduction and factor extraction processes were conducted to check whether the items were distributed according to the a priori factors.

At the evaluation stage, confirmatory factor analysis was conducted with 316 pre-service teachers studying at various teaching programmes and grade levels. Upon gathering the data, dimensionality, validity, and reliability tests were conducted to evaluate the five-factor scale.

Study Group

The study for scale development was conducted with the involvement of 847 pre-service teachers (531 EFA and 316 CFA) from different classes and departments of the Faculty of Education. EFA and CFA data were collected from different groups of students at different times. The distribution of EFA and CFA participants by gender, programme and year is shown in Table 2.

Table 2. *The distribution of EFA and CFA participants.*

	EFA		CFA	
	n	%	*n*	%
Gender				
Male	250	47.10	155	49.05
Female	281	52.90	161	50.95
Grade Level				
2nd Grade	216	40.68	106	33.54
3rd Grade	143	26.93	129	40.82
4th Grade	172	32.39	81	25.63
Department				
Primary Education	146	27.50	55	17.41
Foreign Languages	111	20.90	54	17.09
Psych. Counselling and Guidance	72	13.56	43	13.61
Turkish and Social Studies	61	11.49	41	12.97
Special Education	57	10.73	38	12.03
Math. and Science Education	38	7.16	30	9.49
Fine Arts	36	6.78	32	10.13
Computer and Inst. Technologies	10	1.88	23	7.28
Total	531		316	

Table 2 shows that the participants in the EFA and CFA applications are evenly distributed in terms of gender and grade level. Similarly, the distribution across departments is proportional and reflective of the overall population. The scale development data were collected from the classes under the control of the researcher via pen and paper.

The Data Analysis

After the data were collected, they were first transferred to Excel and then to the Jamovi (Version 2.3.28.0) programme and checked for missing and extreme values. Since missing data was less than 5%, no imputation was performed. After removing outliers and non-respondents, the dataset consisted of 519 participants for EFA. For the CFA, no missing data was observed as the data was collected online. No outliers that could affect the validity and reliability of the dataset were detected in the analyses. After controlling for missing and extreme values, the authors analyzed the dataset to see if it met the assumptions of univariate normality and multivariate normality. In the normality analysis, the kurtosis and skewness values of the EFA and CFA data ranged from -2 to +2, verifying normal distribution. (George & Mallery, 2010).

In addition, the multivariate normality of the EFA and CFA data was analyzed using the Mardia test in order to make a more informed decision about the inference method for the CFA. Analysis of the Mardia test results indicates that the data set fails to meet the criteria for multivariate normality EFA (Skewness = 64.51, Kurtosis = 882.17, p < .001) and CFA (Skewness = 124.48, Kurtosis = 883.04, p < .001). Based on the multivariate normality analysis, Principal Axis Factoring (PAF) was chosen as the factor extraction method in EFA, as it does not rely on the assumption of normality. For the rotation method, Promax rotation was preferred due to its suitability for identifying correlated factors. Regarding the CFA, Maximum Likelihood Robust (MLR), which is more robust to the effects of multivariate normality violations, was preferred as the extraction method in the analysis (Yuan & Bentler, 2000). Following the normality analyses, the correlation matrix was inspected to ensure sufficient intercorrelations among items, with multiple coefficients exceeding 0.3, indicating an appropriate relationship for factor analysis. After that, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity were utilized to assess the factorability of the correlation matrix. The analysis results revealed a significant difference between the unit matrix and the observed correlation matrix (χ^2 = 7343.24, df = 351, p < .001). Both the Bartlett's Test of Sphericity and the KMO value (.94) showed that the data are appropriate for factor analysis (Pallant, 2016).

RESULTS

Exploratory Factor Analysis Results

In the EFA stage, PAF was preferred to check whether the items were distributed in line with the theoretically established model. The oblique rotation method "promax," which accounts for the potential correlations between factors, was selected as the rotation technique (Flora *et al.*, 2012). Following the application of the promax rotation, the factor extraction process was initiated.

Numerous techniques exist for ascertaining the number of factors, with the most commonly used being Kaiser's K1 rule, Cattell's scree plot, and Horn's parallel analysis. Kaiser's K1 rule was originally developed for principal component analysis and is considered unsuitable for principal axis factoring (Kılıç, 2022). Given that, comparative studies have further demonstrated that the parallel analysis method yields significantly more reliable results than the K1 rule (Silverstein, 1987; Zwick & Velicer, 1986). Additionally, research indicates that relying solely on the scree plot often leads to underestimating the number of dimensions (Kılıç, 2022).

Given these considerations, the parallel analysis method was selected for factor extraction as it is widely considered the most reliable approach by many researchers (Hayton *et al.*, 2004). Parallel analysis operates on the principle that the "m largest eigenvalues" derived from the sample's reduced correlation matrix should exceed the "m largest expected" eigenvalues calculated from multiple sets of randomly generated data. These random data sets are generated to match the original sample's size and number of variables (Fabrigar *et al.*, 1999, p. 279). In parallel analysis, eigenvalues from actual data are compared with eigenvalues from randomly generated data (simulations). Factors from actual data with eigenvalues exceeding the corresponding eigenvalue (mean or 95th percentile) in random data are preserved (Hayton *et al.*, 2004; Revelle, 2019).

The initial parallel analysis results indicated that a factorization of up to seven factors can be achieved as the eigenvalues from the actual data exceeded the eigenvalues from the randomly generated data for the first seven factors (Figure 2). This finding supports our a priori theoretical framework, which proposed a five-factor structure, as it falls within the maximum number of meaningful factors suggested by the parallel analysis.

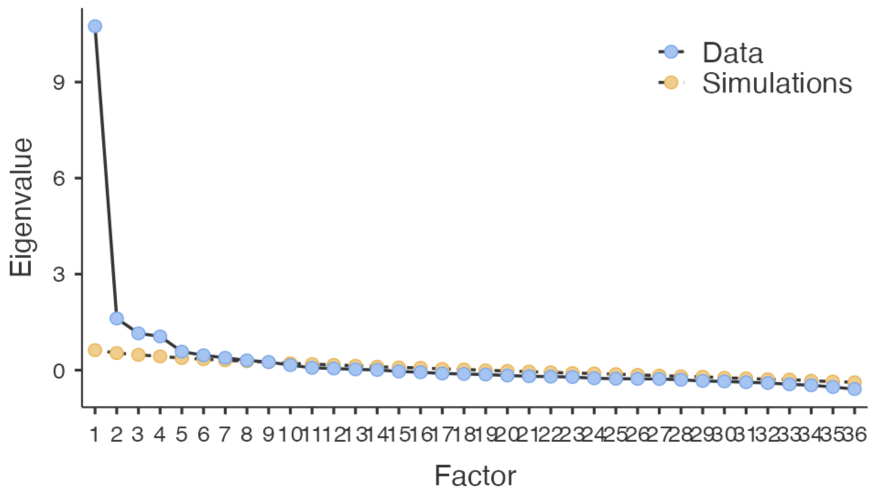

Figure 2. *Scree plot generated by parallel analysis.*

Based on the results of parallel analysis, item loadings on the automatically factorized dimensions were examined, and items loading onto two factors or with factor loadings below .30 were systematically removed. In the initial step, M22, which had a negative loading, and M28, which did not load onto any factor, were excluded from the scale. As the resulting seven-dimensional structure did not align with the predetermined theoretical framework, adjustments were made to ensure consistency. To achieve this, a total of 8 items (M2, M23, M24, M26, M27, M31, M32, M33) that did not load on any factor, or cross-load, or formed a standalone dimension were removed from the scale. Following these modifications, the five-dimensional factor structure, consistent with the theoretical framework and parallel analysis results, is presented in Table 3.

Table 3. *Five-dimensional factor structure.*

Factors	SS Loadings	Eigenvalues	Variance Explained	Total Variance
1	4.52	10.05	16.74	16.74
2	3.18	1.14	11.77	28.50
3	2.78	1.11	10.28	38.78
4	2.06	0.59	7.63	46.40
5	1.65	0.38	6.13	52.53

As Table 3 suggests, the factor structure was created following the theoretical structure and the result of the parallel analysis explains 52.53% of the total variance. Inter-factor correlation coefficients ranged from .47 to .69, indicating acceptable levels of association among the latent constructs. Although parallel analysis supports a five-factor structure, a three-factor solution appears more appropriate according to the K-1 rule. However, the five-factor solution explains 52.52% of total variance, which is above the 50% threshold generally accepted in social sciences. If the fourth and fifth factors were removed, the explained variance would fall below 50%. Therefore, these factors were retained as they contribute meaningful variance and align with the theoretical framework. Literature (Costello & Osborne, 2005; Fabrigar *et al.*, 1999; Kline, 1994) also emphasizes that factors with eigenvalues below 1 can be preserved when they make significant theoretical contributions.

Table 4. *Factor loadings of items.*

		Factors					
		1	2	3	4	5	
		Meta-cognitive Orientation	Affective Orientation	Cognitive Orientation	Quality of Content	Quality of Presentation	
		16.74*	11.77*	10.28*	7.63*	6.13*	
		.93**	.93**	.90**	.90**	.78**	
No	Items	.94***	.92***	.90***	.90***	.78***	Uniqueness
1	M1			.33			.58
2	M3			.51			.72
3	M4			.51			.48
4	M5			.76			.33
5	M6			.67			.31
6	M7			.56			.40
7	M8	.66					.35
8	M9	.70					.39
9	M10	.91					.28
10	M11	.81					.36
11	M12	.53					.50
12	M13	.66					.48
13	M14	.63					.43
14	M15		.56				.44
15	M18		.71				.43
16	M16		.88				.39
17	M17		.88				.27
18	M20		.48				.49
19	M19		.54				.53
20	M34				.42		.63
21	M35				.78		.41
22	M36				.76		.30
23	M37				.79		.49
24	M21					.37	.78
25	M29					.34	.85
26	M25					.51	.69
27	M30					.68	.54

Note. 'Principal axis factoring' extraction method was used in combination with a 'promax' rotation
* Explained variance for factors
** McDonald’s Omega coefficients
*** Composite Reliability coefficients

Consequently, a five-factor solution was preferred. Furthermore, the five-factor solution was found to contribute to the model's significance and consistency. The model fit indices indicated that the five-factor structure contributed significantly to the model’s fit and consistency. Specifically, the RMSEA value was .05, with a 90% confidence interval ranging from .05 to .06, suggesting a good model fit. The Tucker-Lewis Index (TLI) was .93, which indicates an acceptable level of model fit. Additionally, the Bayesian Information Criterion (-860.86) and the chi-square statistic (χ^2 = 522.07, df = 226, p < .001) further supported the model's overall significance. Following the adoption of the five-factor solution, the factor loadings and uniqueness values for each item were examined (see Table 4).

Table 4 shows that the items had factor loadings close to .40 (e.g., M21, M34) or above .40. Additionally, the majority of the items had low uniqueness values (.30–.50), indicating that these items were well-explained by the factors. Overall, the high factor loadings and low uniqueness values offer compelling proof of the scale's structural validity. Consequently, as a result of the EFA, a five-dimensional scale consisting of 27 items was established, comprising the dimensions of 'cognitive orientation,' 'meta-cognitive orientation', 'affective orientation', 'quality of presentation', and 'quality of content'. The findings indicate that the hypothesized five-dimensional structure was successfully formed.

Confirmatory Factor Analysis Results

The scale development phase was followed by the scale evaluation phase. This phase involved 316 pre-service teachers from different teacher education programmes and year levels within the Faculty of Education. The results of the Mardia test showed a violation of multivariate normality, even though the data satisfied the requirements of univariate normality. Therefore, the Maximum Likelihood Robust (MLR) estimation method was used, which is more robust to the effects of multivariate normality violations (Yuan & Bentler, 2000).

The fit indices for the five-factor model successfully converged after 57 iterations with 316 observations and 85 free parameters. In reporting the fit indices, scaled values were used, which are more robust to sample size and normality violations and provide more accurate fit results as shown in Table 5 (Sattora & Bentler, 2001; Yuan & Bentler, 1998). These scaled fit indices are particularly well suited to the challenges posed by non-normality and sample size variation and provide a more reliable assessment of model fit.

Table 5. *Fit indices obtained by CFA.*

Fit Indices	Good Fit	Acceptable Fit	Value	Decision	References
χ^2/df	≤ 2	≤ 3	1.86	Good Fit	Schumacker & Lomax, 2004; Tabachnick & Fidell, 2007
RMSEA	≤ .05	.05–.08	.05	Good Fit	MacCallum *et al*., 1996; Hu & Bentler, 1999
SRMR	≤ .05	≤ .08	.05	Good Fit	Hu & Bentler, 1999
NFI	≥ .95	.90–.95	.90	Acceptable Fit	Byrne, 1994 ; Hu & Bentler, 1999
TLI/NNFI	≥ .95	.90–.95	.94	Acceptable Fit	Hu & Bentler, 1999; Schumacker & Lomax, 2004
CFI/RNI	≥ .95	.90–.95	.95	Acceptable Fit	Kline, 2016
GFI	≥ .95	≥ .90	.94	Good Fit	Jöreskog & Sörbom, 1984
AGFI	≥ .90	≥ .85–.90	.92	Good Fit	Byrne, 1994
IFI	≥ .95	.90–.95	.95	Good Fit	Hu & Bentler, 1999

As shown in Table 5, the scale demonstrates good fit indices, with TLI, NFI, and NNFI values being within acceptable ranges. Following the evaluation of the fit indices, item factor loadings and inter-factor correlation values were analyzed. In the initial results, item M2 was excluded from the scale due to its low factor load (β = -0.12). Subsequently, within the refined structure, it was found that item M8 exhibited strong correlations (r = .58–.70) with the items under the 'Cognitive Orientation dimension', namely M1, M6, M3, M4, and M5. Item M8 was removed from the scale due to the excessive correlation (r = .90) between the BY and ÜBY dimensions.

This adjustment increased the distinctiveness of the dimensions, thereby improving construct validity and aligning the scale more closely with the theoretical framework. Following the removal of these two items, the item factor loadings, error variances, and inter-dimensional correlation values are presented in Figure 3.

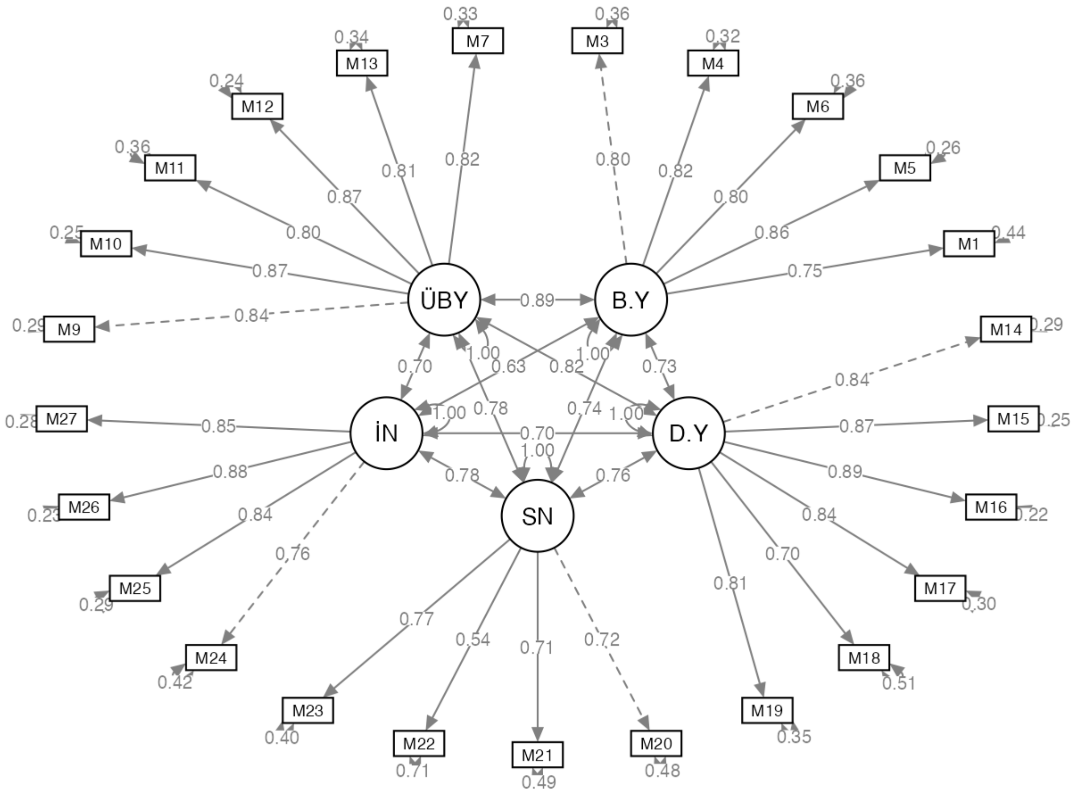

Figure 3. *Path analysis obtained by CFA.*

Figure 3 shows that the standardized factor loadings are above .50, indicating moderate to high levels of acceptability (Tabachnick & Fidell, 2013). This finding supports the convergent validity and structural validity of the scale. In addition, the standardized error variances for the items ranged between .22 and .51. An analysis of the correlations between the latent variables revealed coefficients ranging from .63 to .89. While the value of .89 indicates a strong relationship between the latent variables, the fact that the correlation coefficients are below .90 indicates that the factors are distinct, which supports the validity of discriminant analysis. (Kline, 2011).

Finally, the reliability coefficients and the Average Variance Extracted (AVE) values for the CFA solution were examined. The reliability and validity of the instrument were evaluated using Cronbach's alpha (α), McDonald's Omega (ω), and Average Variance Extracted (AVE) values. The recommended threshold of .70 was exceeded by the Cronbach's alpha values, which varied from .78 to .93, indicating high internal consistency for all dimensions (Nunnally & Bernstein, 1994). For instance, the 'affective orientation' dimension demonstrated strong reliability with an alpha value of .93. Similarly, the Omega coefficients for McDonald's, which range from .77 to .93, further confirmed the high internal consistency of the instrument (Zinbarg *et al.*, 2005). The 'cognitive orientation' dimension, for example, showed a robust reliability with both alpha and omega values of .90. The Cronbach's alpha values obtained for the dimensions estimate the internal consistency of the scale, and the results (α > .70) indicate that the constructs are highly reliable (Nunnally & Bernstein, 1994). Similarly, the McDonald's Omega coefficients indicate that the instrument has high internal consistency and can be considered reliable (Zinbarg *et al.*, 2005). The AVE values provide critical evidence of both discriminant validity and the appropriateness of summing the scores for the factors. The AVE measures the extent to which a factor is well represented by its associated indicators, with values above 0.50 considered acceptable. However, researchers argue that in cases where other reliability measures are high, AVE values between .40 and .50 can also be considered acceptable (Fornell & Larcker, 1981; Hair *et al.*, 2010). When AVE values fall below .50, an alternative, more precise and robust method for assessing the discriminant validity between factors is the Heterotrait-Monotrait Ratio (HTMT) test (Henseler *et al.*, 2015). For the dimension 'quality of presentation' the AVE value was found to be .47. Therefore, the HTMT values for this dimension were examined and were found to be in the range of .73 to .80. These scores provide significant proof for the discriminant validity of the scale and the appropriateness of obtaining a total score.

DISCUSSION and CONCLUSION

The objective of this study was to develop a valid and reliable scale for assessing and evaluating the quality of feedback comprehensively from the perceptions of teacher candidates. The study involved 847 pre-service teachers from different grade levels and teacher education programmes within the Faculty of Education at a public university. In the context of the study, the five-dimensional structure established by EFA, namely cognitive orientation, meta-cognitive orientation, affective orientation, quality of presentation, and quality of content, is in line with the theoretical perspective proposed by Narciss and Huth (2004). This a priori established structure was further confirmed by CFA, resulting in a reliable and valid scale. The fit indices showed that the model has an overall good fit, while the reliability indices indicate that the model is consistent and reliable. Furthermore, the obtained reliability values, CFI, RMSEA, other fit indices, AVE, HTMT values, and the theoretical basis provide significant evidence and support for obtaining a total score from the scale (Hair *et al.*, 2010; Kline, 2011). Examples of items for each dimension and the final structure of the scale are provided in English translation (see Appendix 2) and the original Turkish version (see Appendix 3).

Practical Implications

The developed scale provides a reliable and valid tool for assessing key dimensions of feedback quality, including cognitive orientation, meta-cognitive orientation, affective orientation, quality of presentation, and quality of content. The scale's practical implications extend to various areas. First of all, the scale provides educators with a systematic and thorough method to assess the quality of feedback received by students. By recognizing strengths and weaknesses across various dimensions, educators can tailor their feedback strategies to improve learning outcomes. Secondly, researchers and educators can utilize the scale to evaluate the effectiveness of their feedback in practices both in experimental and applied educational settings. Additionally, this tool can be utilized in initial teacher education to raise awareness about feedback practices and their role in learning. Furthermore, instructional designers and curriculum specialists can utilize the tool for designing effective feedback practices and thereby effective instructional designs with evidence-based principles. Last but not least, the scale can serve as a framework for institutions aiming to standardize and enhance feedback practices as part of their quality assurance measures in education.

Limitations and Recommendations

Although the research process has been carried out using rigorous scientific methods and the most up-to-date sources, there are some limitations associated with the research. Firstly, the study was conducted on a limited sample of pre-service teachers from the Faculty of Education of a state university in Turkey. Despite the inclusion of participants from a wide range of teaching programmes (Classroom Teaching, Foreign Languages, Psychological Counselling and Guidance, Turkish and Social Sciences, Special Education, Mathematics and Science Education, Fine Arts, and Computer Education and Instructional Technology), the distribution among the programmes seems uneven. Although attention has been paid to the proportional distribution, the distribution among the programmes is not exactly equal as the departments of Primary

Education (24.1%) and Foreign Languages (19.7%) account for almost half of the sample. This programme imbalance may have influenced our findings as feedback expectations and experiences in different disciplinary areas may differ significantly.

Secondly, institutional factors also limit our findings that the feedback culture and practices at this particular public university may differ significantly from private institutions or universities in different regions with different pedagogical approaches, lecturer-student ratios and resources available to provide detailed feedback. The particular curriculum structure and assessment methods used in this teacher education programme also shape students' exposure to and expectations of feedback quality. This may limit the generalizability of the scale to pre-service teachers in other universities or other disciplines.

Additionally, the reliability and validity of the scale were evaluated only within the scope of this study, and its stability over time (e.g., test-retest reliability) was not analyzed. Examining the consistency of the scale over time constitutes an important limitation in terms of evaluating whether it gives similar results in different periods. Therefore, conducting test-retest analyses in future studies would be an important step to evaluate the stability of the measurement tool. Moreover, while the current study established robust construct validity through exploratory and confirmatory factor analyses, criterion validity studies were not conducted. Future research should investigate the scale's relationship with external criteria, such as student learning outcomes, academic performance indicators, or other established feedback assessment measures. Exploring these external relationships would provide additional evidence of the scale's predictive and concurrent validity, thereby enhancing its overall utility and generalizability across different educational contexts.

Finally, reliance on self-reported data is a limitation, as individuals' perceptions and memories may not accurately reflect their actual behavior or objectively observable practices (Akbulut *et al.*, 2023; Erdem Kara & Akbulut, 2025; Zhou & Winne, 2012). Social desirability bias may have influenced participants' responses, especially with regard to their interactions with feedback. The following recommendations for further research can be made in light of the study's limitations;

- The scale's reliability and validity should be examined in various settings in order to improve the study's generalizability by working with individuals from different universities, disciplines other than teaching, or different age groups.
- The scale can be adapted and applied to different cultural contexts. This will allow the validity of the scale to be assessed at an international level and cross-cultural comparisons to be made.
- By assessing the consistency of the scale over time (test-retest reliability), it should be examined whether the measurement tool gives similar results over different time periods.
- In addition to construct validity, other types of validity, such as criterion validity and predictive validity, can be evaluated to offer a more extensive evaluation of the scale.

Acknowledgments

This study is based on a part of the first author's doctoral dissertation.

Declaration of Conflicting Interests and Ethics

The authors declare no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in IJATE belongs to the authors. **Ethics Committee Number**: Anadolu University Ethics Committee, 24.06.2022-329898.

Contribution of Authors

Emrullah Esen: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Software, and Writing-original draft. **Oktay Cem Adıgüzel**: Methodology, Supervision, Validation, Reviewing, and Editing. **Derya Atik Kara:** Methodology, Supervision, Validation, Reviewing, and Editing.

Orcid

Emrullah Esen https://orcid.org/0000-0002-7301-4986
Oktay Cem Adıgüzel https://orcid.org/0000-0002-7985-4871
Derya Atik Kara https://orcid.org/0000-0002-6890-030X

REFERENCES

Akbulut, Y., Saykılı, A., Öztürk, A., & Bozkurt, A. (2023). What if it's all an illusion? To what extent can we rely on self-reported data in open, online, and distance education systems? *The International Review of Research in Open and Distributed Learning, 24*(3), 1-17. https://doi.org/10.19173/irrodl.v24i3.7321

Akkuzu, N., & Uyulgan, M.A. (2014). Toward making the invisible visible using a scale: Prospective teachers' thoughts and affective reactions to feedback. *Irish Educational Studies*, *33*(3), 287–305. https://doi.org/10.1080/03323315.2014.923184

Baydas Onlu, O., Abdusselam, M.S., & Yilmaz, R.M. (2022). Students' perception of instructional feedback scale: Validity and reliability study. *Contemporary Educational Technology, 14*(3), Article ep368. https://doi.org/10.30935/cedtech/11811

Beydogan, H.Ö. (2016). Feedback correction perception scale for teacher candidates. *Ahi Evran University Journal of Kırşehir Education Faculty*, *17*(2), 297–314.

Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgar-Quiñonez, H., & Young, S.L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Brookhart, S.M. (2008). *How to give effective feedback to your students.* Association for Supervision and Curriculum Development.

Byrne, B.M. (1994). *Structural equation modeling with EQS and EQS/Windows*. Sage.

Costello, A.B., & Osborne, J.W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation*, *10*(1), Article 7. https://doi.org/10.7275/jyj1-4868

Deeley, S.J., Fischbacher-Smith, M., Karadzhov, D., & Koristashevskaya, E. (2019). Exploring the 'wicked' problem of student dissatisfaction with assessment and feedback in higher education. *Higher Education Pedagogies*, *4*(1), 385-405. https://doi.org/10.1080/23752696.2019.1644659

Deci, E.L., & Ryan, R.M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media. https://doi.org/10.1007/978-1-4899-2271-7

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299. https://doi.org/10.1037/1082-989X.4.3.272

Flora, D.B., LaBrish, C., & Chalmers, R.P. (2012). Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis. *Frontiers in Psychology*, *3*, Article 55. https://doi.org/10.3389/fpsyg.2012.00055

Fornell, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. https://doi.org/10.1177/002224378101800104

George, D., & Mallery, M. (2010). *SPSS for Windows step by step: A simple guide and reference, 17.0 update* (10th ed.) Pearson.

Glazzard, J., & Stones, S. (2019). Student perceptions of feedback in higher education. International *Journal of Learning, Teaching and Educational Research*, *18*(11), 38-52. https://doi.org/10.26803/ijlter.18.11.3

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010) *Multivariate Data Analysis* (7th ed.). Pearson.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hayton, J.C., Allen, D.G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*(2), 191–205. https://doi.org/10.1177/1094428104263675

Henderson, M., Ryan, T., & Phillips, M. (2019). The challenges of feedback in higher education. *Assessment & Evaluation in Higher Education, 44*(8), 1237-1252. https://doi.org/10.1080/02602938.2019.1599815

Henseler, J., Ringle, C.M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Holstun, V., Rigsbee, N., & Bohecker, L. (2021). Encouragement is not enough: perceptions and attitudes towards corrective feedback and their relationship to self-efficacy. *Teaching and Supervision in Counseling, 3*(3), Article 2. https://doi.org/10.7290/tsc030302

Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jöreskog, K.G., & Sörbom, D. (1984). *LISREL VI user's guide* (3rd ed.). Scientific Software.

Erdem Kara, B., & Akbulut, Y. (2025). Prevalence and psychometric implications of careless responses in an online student survey. *Journal of Psychoeducational Assessment, 43*(5), 467-487. https://doi.org/10.1177/07342829251328132

Kara, F.M., Kazak, F.Z., & Asci, F.H. (2018). Perceived Teacher Feedback Scale: The validity and reliability study. *Hacettepe Journal of Sport Sciences, 29*(2), 79-86. https://doi.org/10.17644/sbd.306544

Kartol, A., & Arslan, N. (2021). Turkish version of the Feedback Orientation Scale: Investigation of psychometric properties. *International Journal of Turkish Literature Culture Education, 10*(1), 321–329. https://doi.org/10.7884/teke.5129

Kılıç, A.F. (2022). Deciding the number of dimensions in explanatory factor analysis: A brief overview of the methods. *Pamukkale University Journal of Social Sciences Institute, 51*(1), 305–318.

Kline, P. (1994). *An easy guide to factor analysis*. Routledge.

Kline, R.B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Kline, R.B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.

Koka, A., & Hein, V. (2003). Perceptions of teacher's feedback and learning environment as predictors of intrinsic motivation in physical education. *Psychology of Sport and Exercise, 4*(4), 333–346. https://doi.org/10.1016/S1469-0292(02)00012-2

Kumral, O., & Saracaloğlu, A.S. (2011). An evaluation of elementary school teacher's teaching profession courses programme via educational criticism model. *Education Sciences, 6*(1), 106-118.

Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563-575.

London, M., & Smither, J.W. (2002). Feedback orientation, feedback culture, and the longitudinal performance management process. *Human Resource Management Review, 12*(1), 81-100.

MacCallum, R.C., Browne, M.W., & Sugawara, H.M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130-149. https://doi.org/10.1037/1082-989X.1.2.130

Mulliner, E., & Tucker, M. (2017). Feedback on feedback practice: perceptions of students and academics. *Assessment & Evaluation in Higher Education, 42*(2), 266-288. https://doi.org/10.1080/02602938.2015.1103365

Narciss, S., & Huth, K. (2004). How to design informative feedback for multimedia learning. In Niegemann, R. Brünken, & D. Leutner (Eds.), *Instructional design for multimedia learning*. Waxmann.

Nugraheny, E., Claramita, M., Rahayu, G.R., & Kumara, A. (2016). Feedback in the nonshifting context of the midwifery clinical education in Indonesia: A mixed methods study. *Iranian Journal of Nursing and Midwifery Research, 21*(6), 628. https://doi.org/10.4103/1735-9066.197671

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

O'Donovan, B., Rust, C., & Price, M. (2016). A scholarly approach to solving the feedback dilemma in practice. *Assessment & Evaluation in Higher Education, 41*(6), 938-949. https://doi.org/10.1080/02602938.2015.1052774

Ocak, G., & Karafil, B. (2020). Development of Teacher Feedback Use Evaluation Scale. *International Journal of Progressive Education, 16*(1), 287-299. https://doi.org/10.29329/ijpe.2020.228.20

Ostaeyen, S.V., Embo, M., Rotsaert, T., Clercq, O.D., Schellens, T., & Valcke, M. (2023). A qualitative textual analysis of feedback comments in e-portfolios: Quality and alignment with the CanMEDS roles. *Perspectives on Medical Education*, *12*(1). https://doi.org/10.5334/pme.1050

Öntaş, T., & Kaya, B. (2019). Investigation of the opinions of pre-service primary school teachers to give feedback in preparation of teaching materials. *Journal of Milli Eğitim*, *48*(224), 59–73.

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using SPSS program* (6th ed.). McGraw-Hill Education.

Panadero, E., & Lipnevich, A.A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, *35*, Article 100416. https://doi.org/10.1016/j.edurev.2021.100416

Prilop, C.N., Weber, K.E., Prins, F.J., & Kleinknecht, M. (2021). Connecting feedback to self-efficacy: Receiving and providing peer feedback in teacher education. *Studies in Educational Evaluation, 70*, Article 101062. https://doi.org/10.1016/j.stueduc.2021.101062

Revelle, W. (2019). *Psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

Sattora, A., & Bentler, P.M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507–514. https://doi.org/10.1007/BF02296192

Schumacker, R.E., & Lomax, R.G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Erlbaum.

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

Silverstein, A.B. (1987). Note on the parallel analysis criterion for determining the number of common factors or principal components. *Psychological Reports, 61*(2), 351-354. https://doi.org/10.2466/pr0.1987.61.2.351

Şahin, M. (2015). Investigation of prospective teachers' opinions about the feedback activity used in teaching and learning process. *Bolu Avant Izzet Baysal University Journal of Faculty of Education*, *15*(USBES Special Issue I), 247–264.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.

Tabachnick, B.G., & Fidell, L.S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Williams, A. (2024). Delivering effective student feedback in higher education: An evaluation of the challenges and best practice. *International Journal of Research in Education and Science (IJRES), 10*(2), 473-501. https://doi.org/10.46328/ijres.3404

Winstone, N., Nash, R.A., Rowntree, J., & Parker, M. (2017). 'It'd be useful, but i wouldn't use it': barriers to university students' feedback seeking and recipience. *Studies in Higher Education, 42*(11), 2026-2041. https://doi.org/10.1080/03075079.2015.1130032

Yang, L.; Sin, K.; Li, X.; Guo, J. & Lui, M. (2014). Understanding the power of feedback in education: A validation study of the Feedback Orientation Scale (FOS) in classrooms. *The International Journal of Educational and Psychological Assessment, 16*(1), 21-46.

Yuan, K.-H., & Bentler, P.M. (1998). Robust mean and covariance structure analysis. *British Journal of Mathematical and Statistical Psychology, 51*(1), 63-88. https://doi.org/10.1111/j.2044-8317.1998.tb00667.x

Yuan, K.-H., & Bentler, P.M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology, 30*(1), 165-200. https://doi.org/10.1111/0081-1750.00078

Zinbarg, R.E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's alpha, Revelle's beta, and McDonald's omega: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133. https://doi.org/10.1007/s11336-003-0974-7

Zhou, M., & Winne, P.H. (2012). Modeling academic achievement by self-reported versus traced goal orientation. *Learning and Instruction, 22*(6), 413-419. https://doi.org/10.1016/j.learninstruc.2012.03.004

Zwick, W.R., & Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*(3), 432–442. https://doi.org/10.1037/0033-2909.99.3.432

APPENDICES

Appendix 1. *Lawshe's (1975) CVR and CVI for the Quality of Feedback Scale*.

The Perceived Quality of Feedback Scale
Content Validity Ratios and Index

Item Number	Stayed	Not Stayed	Total	CVR	Decision (>= .59)	Item Number	Stayed	Not Stayed	Total	CVR	Decision (> =.59)
1	12	0	12	1	I*	21	12	0	12	1	I
2	9	0	12	1	I	22	12	0	12	1	I
3	12	0	12	1	I	23	12	0	12	1	I
4	12	0	12	1	I	24	12	0	12	1	I
5	12	0	12	1	I	25	12	0	12	1	I
6	12	0	12	1	I	26	11	1	12	0,83	I
7	12	0	12	1	I	27	12	0	12	1	I
8	12	0	12	1	I	28	12	0	12	1	I
9	11	1	12	0,83	I	29	12	0	12	1	I
10	10	2	12	0,66	I	30	12	0	12	1	I
11	11	1	12	0,83	I	31	12	0	12	1	I
12	12	0	12	1	I	32	12	0	12	1	I
13	12	0	12	1	I	33	10	2	12	0,66	I
14	12	0	12	1	I	34	12	0	12	1	I
15	12	0	12	1	I	35	12	0	12	1	I
16	12	0	12	1	I	36	12	0	12	1	I
17	12	0	12	1	I	37	12	0	12	1	I
18	11	1	12	0,83	I	38	5	0	0	1	I
19	11	1	12	0,83	I			CVI= 0,95			

Appendix 2. *The Perceived Quality of Feedback Scale (English Translated Version)*.

The Perceived Quality of Feedback Scale							
In the context of this research, the term "feedback" refers to comments and evaluations provided by instructors regarding your work, your performance or yourself in the courses, either in written or verbal form, including exams (The following items are the continuation of the sentence in quotation marks).			**Never**	**Rarely**	**Sometimes**	**Often**	**Always**
	No	**Items**					
	"I receive feedback that;						
Cognitive Orientation	1	shows the strengths and weaknesses of my work.					
	2	explains the progress I have made.					
	3	improves my implementation skills.					
	4	improves my analysis skills.					
	5	offers guidance to improve my work.					
Meta-cognitive Orientation	6	improves my learning strategies.					
	7	enhances my learning process.					
	8	raises my awareness regarding my learning process.					
	9	improves my research skills.					
	10	enables me to evaluate my learning process.					
	11	develops my critical thinking skills.					
Affective Orientation	12	increases my learning motivation.					
	13	makes me feel valued.					
	14	supports my academic self-confidence.					
	15	encourages my participation in the lesson.					
	16	is emotionally positive even if I give incomplete or incorrect answers.					
	17	appreciates the strong points of my work.					
	"I receive feedback;"						
Quality of Presentation	18	for each of my work.					
	19	face-to-face.					
	20	from my peers.					
	21	immediately following my work.					
Quality of Content	22	with detailed information on my work.					
	23	that is relevant to objectives of the course.					
	24	that is relevant to my learning needs.					
	25	focusing on my deficiencies in the subject.					

Appendix 3. *The Perceived Quality of Feedback Scale (Original Turkish Version).*

GERİ BİLDİRİM NİTELİK ALGISI ÖLÇEĞİ							
*Bu çalışma kapsamında geri bildirimler, yazılı ya da sözlü olarak gerçekleştirdiğiniz çalışmalarınız/ performanslarınız kapsamında (sınavlar dahil) öğretim elemanları tarafından çalışmalarınıza/performanslarınıza ya da size yönelik yapılan yorum ve değerlendirmeleri ifade etmektedir. (Aşağıda yer alan maddeler tırnak içerisindeki cümlenin devamı niteliğindedir.) **"Eğitim fakültesi lisans dersleri kapsamında; …"**			**Hiçbir Zaman**	**Nadiren**	**Ara Sıra**	**Sıklıkla**	**Her zaman**
	No	**Maddeler**					
Bilişsel Yönelim	1	Çalışmalarımın güçlü ve zayıf yönlerini gösteren geri bildirimler verilir.					
	2	Gösterdiğim gelişimi açıklayan geri bildirimler verilir.					
	3	Uygulama becerilerimi geliştirmeye dönük geri bildirimler verilir.					
	4	Analiz becerilerimi geliştirmeye dönük geri bildirimler verilir.					
	5	Çalışmalarımı geliştirmem için yönlendirmeler sunan geri bildirimler verilir.					
Üst-Bilişsel Yönelim	6	Öğrenme stratejilerimi geliştirmeme yönelik geri bildirimler verilir.					
	7	Öğrenme sürecimi etkili kılmaya yönelik geri bildirimler verilir.					
	8	Öğrenme sürecime ilişkin farkındalık kazandırmaya yönelik geri bildirimler verilir.					
	9	Araştırma becerilerimi geliştirmeme yönelik geri bildirimler verilir.					
	10	Öğrenme sürecimi değerlendirmeme yönelik geri bildirimler verilir.					
	11	Düşünme ve sorgulama becerilerimi geliştiren geri bildirimler verilir.					
Duyuşsal Yönelim	12	Öğrenme motivasyonumu arttırmaya yönelik geri bildirimler verilir.					
	13	Kendimi değerli hissettiren geri bildirimler verilir.					
	14	Akademik özgüvenimi desteklemeye yönelik geri bildirimler verilir.					
	15	Derse katılımımı cesaretlendirmeye yönelik geri bildirimler verilir.					
	16	Eksik/yanlış yanıtlar versem de duygusal yönden olumlu geri bildirimler verilir.					
	17	Çalışmalarımın güçlü yönlerini takdir etmeye yönelik geri bildirimler verilir.					
Sunumun Niteliği	18	Her bir çalışmam için geri bildirim verilir.					
	19	Geri bildirimler yüz yüze verilir.					
	20	Akranlarım tarafından geri bildirim verilir.					
	21	Çalışmalarıma yönelik anlık geri bildirim verilir.					
İçeriğin Niteliği	22	Geri bildirimler çalışmalarım hakkında detaylı bilgiler içerir.					
	23	Geri bildirimler dersin amaçları ile uyumludur.					
	24	Geri bildirimlerin içeriği öğrenme ihtiyaçlarımla uyumludur.					
	25	Geri bildirimler konu hakkındaki eksikliklerime yönelir.					