ISSN 1512-1801

UDC

PERCEPTIONS OF COMPETENCY SCALE TOWARDS ARTIFICIAL INTELLIGENCE-BASED FEEDBACK: VALIDITY AND RELIABILITY STUDY

Sedef Çelik Demirci*,[1], Hatice Kumandaş Öztürk[2] Zeynep Karataş[3]

[1]Department of Mathematics and Science Education, Faculty of Education, Artvin Çoruh University,Artvin/Turkey
[2]Department of Educational Sciences, Faculty of Education, Artvin Çoruh University,Artvin/Turkey
[3]Artvin Çoruh University,Artvin/Turkey

*Corresponding author: sedefcelik@artvin.edu.tr

Abstract
This research was conducted to develop a valid and reliable scale to measure middle school students' perceptions of competency regarding feedback received from artificial intelligence. Data were collected from students in the 5th, 6th, 7th, and 8th grades from five middle schools selected through typical case sampling. Preliminary trials of the scale were conducted with data from 405 middle school students, while the final scale was validated with data from 275 students. Data was analysed using SPSS 27.0 and LISREL 8.80. The initial version of the scale was developed as a 20-item scale with a 5-point Likert scale. Exploratory factor analysis resulted in the removal of overlapping items, refining the scale to a final version that includes 18 items. Results from the Exploratory Factor Analysis (EFA) and parallel analysis indicated that the scale has a single-factor structure, accounting for approximately 35% of the total variance. To verify the adequacy of the resulting 18-item unidimensional model, Confirmatory Factor Analysis (CFA) was performed. The CFA results demonstrated that the model exhibited acceptable fit indices. The reliability of the measurement tool was assessed using Cronbach's alpha coefficient, and the findings revealed a high level of internal consistency, with a coefficient of 0.88.In conclusion, it was determined that the "Perception of Competency towards AI-based Feedback" Scale aligns with the one-dimensional theoretical model and falls within acceptable limits, making it a valid and reliable instrument for use in the literature.

Keywords: *artificial intelligence, competency perception, feedback, measurement tool development.*

Introduction

When examining the development of artificial intelligence, it becomes evident that this concept initially emerged as an idea. However, over time, it evolved into a concrete structure through the advancement of various parameters. Artificial intelligence and AI-based studies are currently applied across various fields, including education [1], and continue to evolve rapidly [2]. AI applications have started to be integrated into the education process, influencing everything from teaching methods to the evaluation of various human-specific skills, including cognitive processes. Similar to other fields, the use of artificial intelligence is widespread in subfields such as robotic tools, expert systems, autonomous planning, logistical planning, speech recognition, and language translation [3-6]. Furthermore, it is believed that

ISSN 1512-1801

artificial intelligence can enhance students' learning experiences and improve educational outcomes [7]. In this context, emphasis is placed on integrating artificial intelligence into preschool and primary education to enhance student achievement through this process [8]. This is because artificial intelligence is recognized for its ability to tailor instruction to individual learning styles, which is believed to enhance student engagement and academic outcomes [9].

Current advancements in artificial intelligence are paving the way for radical transformations in education by increasing opportunities for personalized learning [10]. Therefore, AI applications that prioritize personalization in education (smart lesson systems, language processing/translation, virtual classrooms, educational chatbots) are rapidly developing and expanding in scope [7]. For example, SquirrelAI, a China-based company specializing in intelligent adaptive learning systems, is actively working to integrate AI into education to provide personalized educational services to students [1]. In particular, AI's use of advanced analytics to monitor learning speeds, along with deep learning and machine learning, enables this personalization [11]. As a result, AI has the potential to offer students individualized guidance, support, and feedback [12]. It is emphasized that artificial intelligence applications, especially in the field of education, can function like a personalized 'teacher' by providing instant feedback to students' questions, enabling them to identify and correct their mistakes, and thus assisting the teacher as well [13]. It is evident that artificial intelligence supports teachers in this process, and the use of feedback applications in education is also on the rise [14].

Feedback, often referred to as reciprocation, is an interactive communication process that involves the exchange of positive or negative opinions between teacher and learner, thereby promoting learning [15]. Therefore, giving and receiving feedback is a crucial activity at the core of the teaching process for students [16]. In this respect, the ability of artificial intelligence to provide instant feedback is quite important in assessment and monitoring applications in education. Thanks to this, students can find teaching methods suitable for them based on the feedback they receive from artificial intelligence [17]. Accordingly, it is an undeniable fact that students improve and reinforce their learning with the help of feedback received from artificial intelligence, and that the permanence of learning increases [18]. Therefore, the feedback given by teachers in this process has a significant impact on students' future lives [19]. In this context, considering the role of feedback in reinforcing learning and supporting retention, a valid and reliable measurement tool is necessary to determine how students evaluate feedback received from artificial intelligence and their level of benefit from this feedback. Therefore, this study aims to develop a valid and reliable scale to assess students' perceptions of self-efficacy concerning their understanding, evaluation, and utilization of feedback received from artificial intelligence during the learning process. Indeed, artificial intelligence applications are recognized as playing a significant role in various aspects of education and training processes [20]. With the increasing importance of this role, various scale development and adaptation studies related to artificial intelligence are seen in the international and national literature.

Çakan and Akın [21] developed the *Artificial Intelligence Attitude Scale* by drawing on the Computer Attitude Scale originally proposed by Nickell and Pinto (1986), which was later adapted into the Internet Attitude Scale by Durndell and Haag (2002). The scale was revised to ensure cultural relevance within the Turkish context. Additionally, Demir Dülger and Köklü [22] designed a valid and reliable instrument to assess teachers’ and school principals’ perspectives on the use of artificial intelligence in educational settings. Moreover, Aksekili and Kan [17] suggested that the integration of artificial intelligence in education is closely

ISSN 1512-1801

associated with teachers' perceptions of AI. They developed the 'Teachers' Attitude Scale Towards the Use of Artificial Intelligence in Education' and demonstrated its validity and reliability. However, while there are tools in the literature that measure general self-efficacy perception regarding artificial intelligence, tools that directly measure more specific self-efficacy areas in the educational context are less common. Indeed, [23] adapted the "Artificial Intelligence Self-Efficacy Scale," created in a foreign language by Wang and Chuang (2023), into Turkish. However, a literature review revealed a limited number of scales that directly address self-efficacy perception in the context of understanding, evaluating, and incorporating feedback from artificial intelligence tools used in education into the learning process; this study is therefore expected to contribute to filling this gap. It is recognized that artificial intelligence serves as a powerful tool enabling students to conduct self-assessments, set individual learning goals, engage in activities tailored to their learning preferences, and receive timely, personalized feedback [24]. Furthermore, it is acknowledged that AI-driven feedback applications enhance students' learning outcomes [25]. Therefore, the increasing visibility of AI-based applications in the learning process and the impact of feedback on learning outcomes are being considered.

In light of the research and rationale discussed above, a valid and reliable measurement tool is needed to empower middle school students to assess the feedback provided by AI-based applications and to evaluate their competence in using these feedback tools. In this context, the study aims to develop a scale to measure individuals' perceptions of their competence in using AI-based feedback and to examine the psychometric properties of the developed scale.

Method

This study, as a scale development project, was conducted using a survey model within a quantitative research framework. The research followed the steps of the scale development process, including creating the item pool, examining construct validity, and conducting reliability analyses.

Study Group

The research was conducted with 5th–8th grade students from five middle schools in northeastern Turkey during the 2024–2025 academic year. The study group was chosen using typical case sampling, which is a form of purposive sampling. Typical cases encompass situations that are usual and are considered average within society [26]. In this context, a typical case sampling for a study on schools might be limited to selecting a few schools in the city centre that can relatively reflect the general population [27]. Therefore, this study was conducted with middle schools representing the region and students in 5th, 6th, 7th, and 8th grades. The study group included a total of 680 middle school students. Data were collected in two stages. Data collected from 405 students in the first stage were used for EFA. Data collected from 275 students in the second stage were used for CFA. In scale studies, the common view is that 100 people are considered weak, 200 are considered average, 300 are considered good, 500 are considered very good, and 1000 are considered excellent (Comrey & Lee, 1992, cited in [28]). Considering these criteria, it is accepted that the study group included in the research has reached a sufficient sample size for scale development.

Measurement Tool Development Process

In this research, a literature review was conducted to assess students' perceptions of their competence in relation to AI-based feedback. Studies containing the keywords "Scale Development" and "Artificial Intelligence," "Scale Development" and "Feedback," and "Competence Perception" were examined in the YOK National Thesis Centre, Google Scholar, and ERIC databases. Subsequently, opinions were also gathered from students on

ISSN 1512-1801

this topic. Based on the reviews, a pool of 16 items was initially created to determine students' perceptions of competence regarding AI-based feedback. Expert opinions were sought to ensure the content validity of the items, and a preliminary interview process was conducted with the students. The items were presented to a measurement and evaluation expert and two field experts. Based on the experts' evaluations, it was decided that three items needed revision.

Additionally, four more items were added to the scale based on the experts' opinions. This 20-item scale, later revised and restructured, was presented to a Turkish language expert and a Turkish teacher for evaluation in terms of language and expression competence. The assessments from the experts concluded that the items were understandable, with only some items requiring attention to spelling rules. To better assess the suitability of the items for the purpose and to test their comprehensibility, pilot interviews were conducted with three students from different middle school levels. The interviews revealed that the items were appropriate and understandable. However, following factor analysis, it was decided to remove two items. The final version of the scale items is presented in Appendix 1.

Data Collection

Data were gathered using the *AI-Based Feedback Competency Perception Scale* developed by the researcher. The instrument is a five-point Likert-type scale comprising 20 positively worded items along with four demographic variables. Item responses range from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*), with higher scores indicating stronger agreement.

Data Analysis

Measurement instruments employed in research are required to demonstrate adequate reliability (Pallant, 2020). One of the methods used to determine the internal consistency of the applied scales is Cronbach's Alpha coefficient [29]. Then, the sample size and normality assumptions were tested.

Researchers developing scales often begin their work with a large number of individual scale items. They can improve or reduce these items using factor analysis techniques to create a smaller, more consistent subscale [29]. Two main approaches exist in the literature for performing this analysis: EFA and CFA [29]. Therefore, to evaluate the construct validity of the developed scale, EFA was performed using SPSS 27, and CFA was performed using LISREL 8.80 with the obtained data.

Factor analysis is a multivariate statistical method that aims to obtain new and meaningful variables by grouping interrelated variables [30]. This method plays a significant role, especially in the scale development process. Factor analysis is used to evaluate the construct validity of the scale and determine the sub-dimensions of the concept measured. Analysing the relationship between the items in the scale and the overall structure helps clarify the conceptual framework. In this respect, factor analysis is a fundamental tool in the scale creation process .

Validity and reliability constitute essential stages in the scale development process. Validity refers to the degree to which a measurement instrument accurately captures the construct it is designed to measure, without interference from unrelated factors [32]. In other words, it reflects the extent to which a tool measures the intended variable rather than other attributes. Reliability, by contrast, concerns the amount of measurement error present in the results and indicates the consistency of measurements obtained under identical conditions [33-34].

ISSN 1512-1801

One of the most frequently employed indicators of reliability is Cronbach's alpha coefficient, which assesses the internal consistency of a scale. Validity can be examined through several approaches, including construct validity, content validity, and criterion-related validity [35]. In the present study, Exploratory Factor Analysis (EFA) was initially conducted to identify the underlying factor structure of the scale, after which reliability analyses were performed.

The calculations related to the assumptions required prior to conducting the analyses are presented below.

Normality and linearity:
Before starting the analysis, the normality and linearity values of the dataset were examined using graphs. Figure 1 shows the histogram and P-P plot of the collected data for CFA.

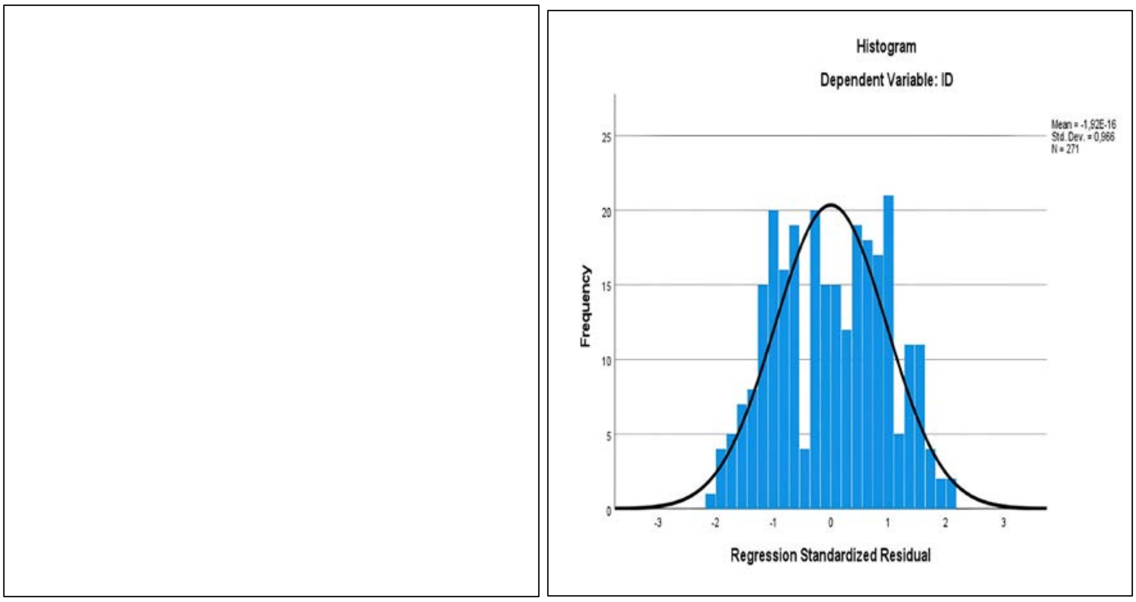

Figure 1. Normality and linearity graphs

Sample size:
The sample size should be greater than 200 in CFA [36-37]. As this value increases from 200 to 500, the suitability in terms of evaluation criteria and the probability of the model being accepted increase. In the second phase of the study, data were collected from 275 individuals for CFA. In this respect, it can be considered that the sample size assumption for the analysis was met.

Examining the missing data:
The dataset was first examined to determine whether the distribution of missing data was random. Analysis of the missing data revealed that, except for four individuals, the missing data were randomly distributed. Therefore, no assignment was made. However, the data of those four individuals were removed from the dataset.

The presence of extreme values:
Standard z-values were calculated for each item in the dataset, and since the obtained z-values were in the range of (-3) to (+3), none of the data were excluded. CFA was continued with data from 271 individuals.

ISSN 1512-1801

Research and Publication Ethics

This study was conducted in accordance with the principles outlined in the *Guidelines for Scientific Research and Publication Ethics in Higher Education Institutions*, and none of the practices identified as violations of scientific research and publication ethics in the second section of these guidelines were undertaken. In addition, ethical approval was obtained from the Ethics Committee of Artvin Çoruh University (Decision No. E-19227540-302.08.01-175874, dated March 28, 2025).

Findings

Findings regarding construct validity

EFA was used to determine the factor structure of the scale and to analyse the relationships of the scale items with these factors [31]. First, the significance of the KMO and Bartlett tests was checked before conducting the factor analysis. For factor analysis, a significant Bartlett test [29] and a KMO index greater than 0.60 are expected [31]. Table 1 shows the KMO and Bartlett test results obtained from the first factor analysis.

Table 1. *KMO-Bartlett Test*

KMO Sampling Adequacy		**0.921**
Bartlett Test	Chi-Square	2186.653
	Df	190
	P	0.000

The analysis revealed that the KMO value exceeded the critical value (KMO = 0.92), and the p-value ($p < 0.01$) was statistically significant, as indicated by the Bartlett test results. This shows that the sample size is sufficient and a multivariate normal distribution has been identified. When determining the number of factors, the eigenvalue criterion, the explained variance ratio, and the scree test graph are considered.

The total variance explained in Table 2 and the scree plot graph given in Figure 2 were examined as a result of the analysis, and the number of factors was determined.

Table 2. *Total variance explained*

	Eigenvalue		
Components	Total	% Variance	Cumulative %
1	6.491	32.453	32.453
2	1.269	6.345	38.798
3	1.144	5.718	44.516
4	.986		
5	.912		
6	.864		
7	.828		
8	.793		
9	.766		
10	.712		
11	.673		
12	.641		
13	.608		
14	.558		
15	.555		
16	.511		

ISSN 1512-1801

17	.461
18	.457
19	.417
20	.354

Table 2 shows that, in the 20-item measurement instrument, three factors are significant and together explain almost 45% of the total variance, as well as the variance related to the scale.

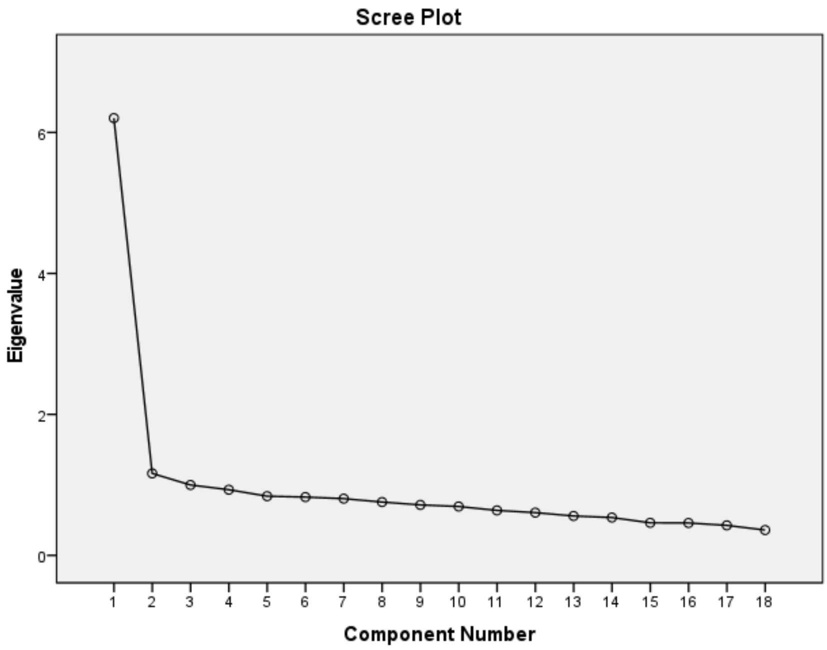

Figure 2. *Line Graph Showing Eigenvector Magnitude*

Figure 2 shows that the measurement instrument appears to be one-dimensional. After this stage of the analysis, factor loadings were examined, and overlapping items were identified to finalize the measurement instrument. Based on the factor loadings, items 2 and 5 were identified as redundant. After removing these items and performing varimax rotation, the factor loadings were summarized in the table.

Table 3. *Components matrix and factor loading values*

	Components
Items	1
M7	.665
M13	.663
M6	.639
M10	.628
M18	.627
M11	.627
M8	.621
M12	.609
M3	.592
M9	.578
M4	.577
M16	.561
M20	.556
M1	.545

ISSN 1512-1801

M17	.530
M14	.507
M15	.501
M19	.500

Table 3 shows that the factor loadings of the 18 items ranged from 0.66 to 0.50. After rotation, the variance explained by this factor alone was identified as approximately 35%. "It is stated that the explained variance ratio should be at least 30% in one-dimensional structures" (Büyüköztürk, 2020, cited in [38]). The final versions of the scale items are given in Table 3. Following this stage, CFA was conducted on the data collected using the revised instrument to validate its structure.

According to the results of Parallel Analysis, one of the methods used to determine the number of factors, the eigenvalues of the real data were compared with those of the simulated and resampled data. The eigenvalue of the real data remained above these values only in the first factor. From the second factor onwards, the curves overlapped, or the real data remained lower. This result indicates that the scale supports a single-factor structure.

Findings regarding the reliability of the measurement instrument

To evaluate the reliability of the 18-item measurement instrument, Cronbach's alpha coefficient was calculated. The results indicated that the unidimensional scale yielded a Cronbach's alpha value of 0.887 ($p < 0.01$), suggesting that the instrument demonstrates an acceptable level of internal consistency.

Confirmatory Factor Analysis (CFA)

CFA results

The values obtained from CFA were interpreted by calculating the path diagram and goodness-of-fit measures. The analysis results are given in Table 4 and Figure 3.

Table 4. *Goodness-of-Fit Values Obtained from CFA Results*

Scale	x^2	sd	x^2/sd	p	AGFI	CFI	NNFI	SRMR	RMSEA
18 items	219.25	135	1,62	0.00	0.90	.0.99	0.99	0.04	0.05

According to Table 4, the ratio of the obtained chi-square value (x^2) to the degrees of freedom (sd) is found to be (x^2/sd=1.62). The acceptable value of this ratio should be $x^2/df \leq 5$ (Kline, 2005). In this respect, the model is considered to have a good fit with the data. When the goodness-of-fit indices are examined, an RMSEA value of 0.05, which is considered quite good, was obtained. An RMSEA value less than 0.05 indicates an excellent fit [39]. When the other fit indices (AGFI, CFI, NNFI and SRMR) are examined, the obtained values are at an acceptable level. Figure 3 shows the representation of the scale's structure, which was validated as a result of the analysis.

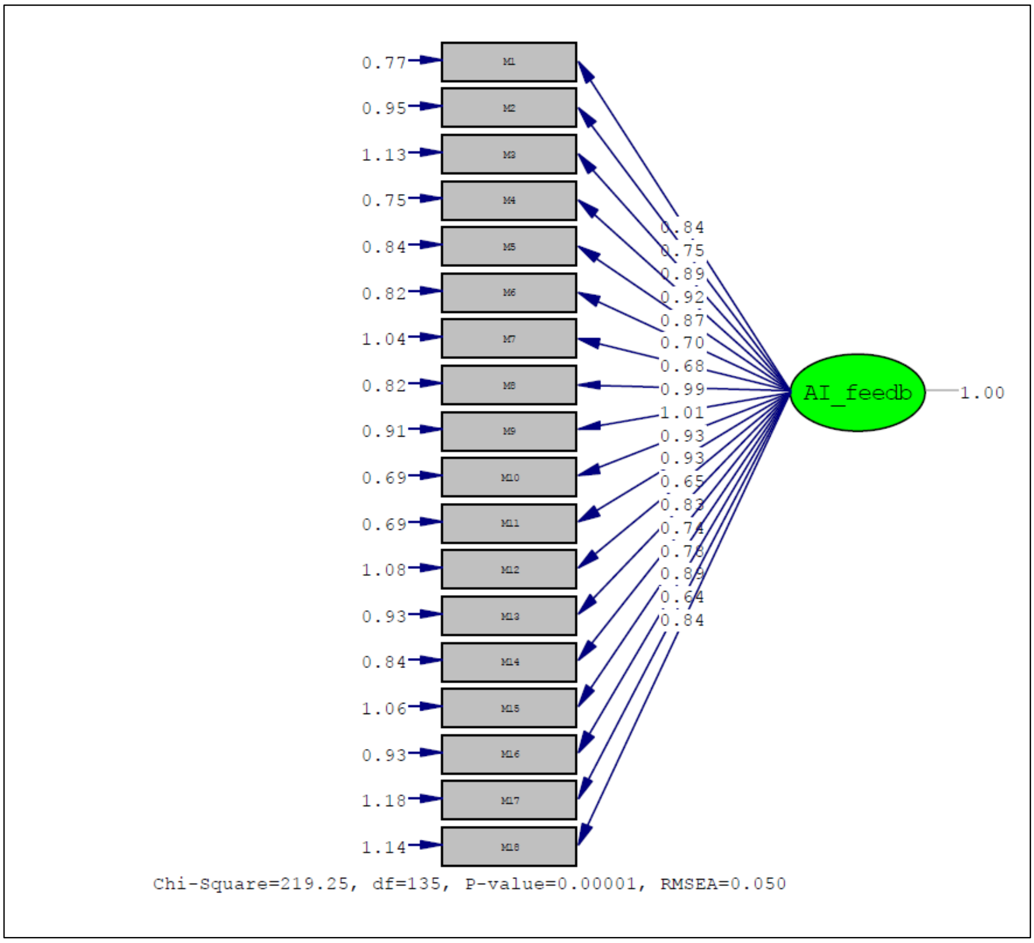

Figure 3. CFA Diagram

Discussion, Conclusion and Recommendations

This study aimed to develop a scale measuring perceived competence toward AI-based feedback and to examine its psychometric properties. The results of the validity analyses indicated that the data met acceptable criteria. Exploratory Factor Analysis (EFA) revealed a unidimensional structure explaining approximately 35% of the total variance, which was subsequently confirmed through Confirmatory Factor Analysis (CFA). Reliability analysis demonstrated satisfactory internal consistency, with a Cronbach's alpha coefficient of 0.88 [40]. Overall, the findings suggest that the *Perceived Competence toward AI-Based Feedback* scale is a valid and reliable measurement instrument.

Ocak and Karanfil [41] developed the *Teacher Feedback Evaluation Scale* to enable students to assess the feedback they receive from their teachers. Their findings indicated that the 23-item instrument has a three-factor structure and demonstrates an adequate level of reliability. In contrast, the present study developed an 18-item scale designed to assess middle school students' perceived competence regarding AI-generated feedback. The results revealed a unidimensional factor structure and satisfactory internal consistency. The observed difference in factor structures suggests that dimensionality may vary depending on both the source of feedback (teacher versus artificial intelligence) and the focus of the measured construct (feedback evaluation versus perceived competence).

Similarly, Morales-García et al. [42] adapted a scale for use in the context of artificial intelligence based on the six-item GSE-6, a shortened version of the ten-item General Self-Efficacy Scale (GSE). Their findings demonstrated a unidimensional structure with a high internal consistency coefficient ($\alpha = 0.91$). The reliability coefficient obtained in the current

ISSN 1512-1801

study (α = 0.88) aligns with the results reported by Morales-García et al. [42], indicating that self-efficacy–related constructs can be measured consistently within AI-oriented assessment frameworks.

In educational settings, feedback provided by teachers plays a critical role in students' academic achievement as well as their social development. Parallel to technological advancements observed in fields such as communication, healthcare, and transportation, significant progress has also been made in the integration of artificial intelligence into education, particularly in recent years [2]. Consequently, the feedback generated by AI systems—now increasingly embedded in educational processes—has gained importance, much like teacher-provided feedback, in shaping learners' experiences and outcomes.

A review of the literature reveals that with the advancement of technology, artificial intelligence has started to be utilized in various fields, including education, and numerous studies have been conducted on this topic [17]. However, it has been found that there is no competency perception scale available for students to evaluate the feedback they receive from artificial intelligence, highlighting a gap in the literature on this topic. This scale will assist researchers studying the evaluation of feedback received from artificial intelligence, and will provide them with ideas on this subject.

This scale, which measures students' perceptions of their competence in evaluating feedback received from artificial intelligence-an important part of the educational process- was developed for middle school students in this study. In this respect, similar measurement tools can be improved for use by students in other educational levels in future studies. Furthermore, it is believed that conducting mixed-methods studies using this tool, alongside qualitative research, will contribute to the field for researchers who wish to conduct in-depth studies on this research topic. Additionally, based on the results obtained with the tool developed in this study, suggestions can be made, and appropriate training can be provided to enhance the effectiveness of artificial intelligence in delivering feedback during learning processes.

References

1. İşler, B., & Kılıç, M. Y. (2021). Eğitimde yapay zekâ kullanımı ve gelişimi. *e-Journal of New Media (eJNM)*, 5(1), 1–11. https://doi.org/10.17932/IAU.EJNM.25480200.2021/ejnm_v5i1001
2. Alan, B., Kırbağ Zengin, F., Keçeci, G. (2024).Yapay Zekâ Tutum Ölçeği (YZTÖ): Geçerlik ve Güvenirlik Çalışması.Artificial Intelligence Attitude Scale (AIAS): Validity and Reliability Study. *Cumhuriyet International Journal of Education, 13*(4): 789-800, DOI: https://dx.doi.org/10.30703/cije.1327949
3. Nilsson, N.J. (2010). *The quest for artificial intelligence: a history of ideas and achievements.* NewYork. NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511819346
4. Russell, S. & Norvig, P. (2010) Artificial intelligence: A Modern Approach. 3rd Edition, Prentice-Hall, Upper Saddle River.
5. Renals, S. & Hain, T. (2010) Speech Recognition In Clark A, Fox C & Lappin S (Ed.), The Handbook of Computational Linguistics and Natural Language Processing (pp. 299-332).
6. Adalı, E. (2017). Yapay Zekâ. *İTÜ Vakfı Dergisi, (75),* 8-13.
7. Gürbüz, T. (2023). Gelecekte öğrenme ve eğitim: Yapay zekâ uygulamaları. *İktisat ve Toplum*, 156, 103–107.
8. Millî Eğitim Bakanlığı, Yenilik ve eğitim teknolojileri genel müdürlüğü. (2024). Eğitimde Yapay Zekâ Uygulamaları Uluslararası Forumu Raporu.

9. Eden, C. A., Chisom, O. N., & Adeniyi, I. S. (2024). Integrating AI in education: Opportunities, challenges, and ethical considerations. *Magna Scientia Advanced Research and Reviews*, *10*(2), 006-013. https://doi.org/10.30574/msarr.2024.10.2.0039
10. Van der Vorst, T., & Jelicic, N. (2019). Artificial intelligence in education:Can AI bring the full potential of personalized learning to education?In 30th European regional ITS conference, Helsinki 2019 205222. Inter-national Telecommunications Society (ITS). https://hdl.handle.net/10419/205222
11. Kengam, J. (2020). Artificial intelligence in education. *Research Gate*, *18*, 1-4. https://doi.org/10.13140/RG.2.2.16375.65445
12. Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, challenges, roles and research issues of artifcial intelligence in education. *Computers and Education: Artifcial Intelligence*, *1*, 100001. https://doi.org/10.1016/j.caeai.2020.100001
13. Demircioğlu, E., Yazıcı , C., & Demir , B. (2024). Yapay zekâ destekli matematik eğitimi: Bir içerik analizi. *International Journal of Social and Humanities Sciences Research (JSHSR)*, *11*(106), 771–785. https://doi.org/10.5281/zenodo.11109449
14. Holmes, W., Bialik, M., and Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning.* The Center for Curriculum Redesign: MA, USA.
15. Cevher, A. Y., Kara, İ., Akbay, M., Oktay, Ö., vd. (2022). Akademik yazımda kullanılan geri bildirim üzerine yapılan çalışmaların incelenmesi. *Eğitim Teknolojisi Kuram Ve Uygulama*, *12(*1), 147-171. https://doi.org/10.17943/etku.955882
16. Phuong Thi Anh Le & Camilla Vásquez (2011) Feedback in teacher education: mentor discourse and intern perceptions, *Teacher Development*, *(15)* 4, 453-470, DOI: 10.1080/13664530.2011.635264
17. Aksekili, E., & Kan, A. (2024). Öğretmenlerin eğitimde yapay zekâ kullanımına yönelik tutum ölçeği Geliştirme: Geçerlik ve güvenirlik çalışması. *21. Yüzyılda Eğitim ve Toplum*, *13*(39), 525-542. https://dergipark.org.tr/en/pub/egitimvetoplum/issue/89829/1471276
18. Güneş, E., & Bülbül, H. İ. (2014). Web ortamında problem temelli öğrenmede farklı geri bildirim stratejilerinin ve internet kullanımına yönelik tutumun öğrenme üzerindeki etkisi. *Türk Eğitim Bilimleri Dergisi*, *12*(2), 109-125.
19. Dönmez, A. (2024). Öğretmenlerin ders sürecinde geri bildirim verme ile ilgili görüşleri. *International Journal of Language Academy*, *12* (4), 236-246. https://doi.org/10.29228/ijla.78976
20. Bulut, M. A., Davarcı, D., Bozdoğan, N. K., Sarpkaya, Y. (2024). Yapay zekanın eğitim üzerindeki etkileri. *Ulusal Eğitim Dergisi*, *4*(3), 976-986. https://doi.org/10.5281/zenodo.10909352
21. Çakan, M., & Akın, A. (2024). Yapay zeka tutum ve değişime hazır olma: İki ölçek uyarlama çalışması. *Econder Uluslararası Akademik Dergi*, *8*(2), 137- 167. https://doi.org/10.35342/econder.1544898
22. Demir Dülger, E., & Köklü, M. (2023). A scale development study to determine the opinions of school administrators and teachers on the use of artificial intelligence in education. ISPEC *International Journal of Social Sciences & Humanities*, *7*(1), 154–174. https://doi.org/10.5281/zenodo.7767140
23. Uyan, U., & Gültekin, S. U. (2024). Yapay zeka öz-yeterlilik ölçeğinin Türkçe'ye uyarlanması: Geçerlilik ve güvenirlik çalışması. *Journal of Research in Business*, *9*(1), 135-148. https://doi.org/10.54452/jrb.1415212
24. Fernández Cuevas, H. (2025). Formative assessment and artificial intelligence: Strategies for human and effective learning. *MLS – Pedagogy, Culture and Innovation (MLSPCI)*, *2*(1), 126–141.
25. Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering*, 1–19. https://doi.org/10.1155/ 2022/5215722

26. Patton, M. Q. (2005). *Qualitative research*. New York: John Wiley & Sons, Ltd.
27. Strauss, A. & Corbin, J. (2014). Basics of qualitative research techniques. New York: Sage Publications.
28. Şahin, M. G., & Boztunç Öztürk, N. (2018). Scale development process in educational field: A content analysis research. *Kastamonu Education Journal, 26*(1), 191-199. https://doi.org/10.24106/kefdergi.375863
29. Pallant, J. (2020). *SPSS kullanma kılavuzu: IBM SPSS ile adım adım veri analizi* (S. Balcı & B. Ahi, Çev.; 4. baskı). Anı Yayıncılık.
30. Büyüköztürk, Ş. (2004). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem A Yayıncılık.
31. Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
32. Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz, Ş. ve Demirel, F. (2014). Bilimsel araştırma yöntemleri (17. Baskı). Ankara: Pegem
33. Ercan, I. and Kan, I. (2004) Reliability and validity in scales. *Uludag Universiy Journal of College of Medicine, 30*, 211-216.
34. Arıkan, R. (2013). *Araştırma yöntem ve teknikleri*. Ankara: Nobel Yayınevi.
35. Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741
36. Hair, J., Anderson, R., Tatham, R. and Black, W. (1998) *Multivariate data analysis*. 5th Edition, Prentice Hall, New Jersey.
37. Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). The Guilford Press.
38. Acar Güvendir, M., Kılıç, A. F., & Özer Özkan, Y. (2025). Ölçek seçiminde arayış: Doğru ölçeği bulma yolculuğu. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi (64),* 338-373. https://doi.org/10.9779/pauefd.1501330
39. Çokluk, O., Şekercioğlu, G., & Büyüköztürk, S. (2010). *Sosyal bilimler için çok degiskenli istatistik SPSS ve LISREL uygulamaları.* Ankara: Pegem A.
40. DeVellis, R. F. (2017). Scale development: Theory and applications (4th ed.). Sage.
41. 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
42. Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B., & Morales-García, M. (2024). Adaptation and psychometric properties of a brief version of the general self-efficacy scale for use with artificial intelligence (GSE-6AI) among university students. Frontiers in Education, 9. https://doi.org/10.3389/feduc.2024.1293437

Appendix-1

PERCEPTIONS OF COMPETENCY SCALE TOWARDS ARTIFICIAL INTELLIGENCE-BASED FEEDBACK

Dear students, this study aims to determine middle school students' evaluations of the feedback from artificial intelligence/AI technologies (ChatGPT, Khan Academy, Duolingo, etc.). Therefore, honestly answering each item on the scale is crucial for the validity and reliability of the study.

Thank you for your participation.

	1	2	3	4	5

ISSN 1512-1801

	Items	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I can evaluate the quality of the feedback I receive from AI (how responsive it is to my needs and how useful it is).					
2	I have no difficulty understanding the feedback I receive from AI.					
3	If I don't understand the feedback I receive from AI, I can ask AI another question.					
4	I can seamlessly manage the process of asking questions and receiving feedback when using AI.					
5	I can enjoy the process of asking questions and receiving feedback when using AI.					
6	I feel quite comfortable asking questions and receiving feedback when using AI.					
7	I feel quite peaceful asking questions and receiving feedback when using AI.					
8	I find the process of asking questions and receiving feedback easy when using AI.					

	Items	1	2	3	4	5
		Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
9	I have no difficulty understanding the syntax used by AI when providing feedback.					
10	I receive feedback from AI that makes it easier for me to learn.					
11	I learn more by receiving feedback from AI.					
12	The feedback I receive from AI increases my motivation.					
13	Thanks to the different types of feedback I receive from AI, I become aware of my shortcomings.					
14	I can understand different aspects of the feedback I receive from AI.					

15	I can take precautions against negative feedback I receive from AI.					
16	I can easily use applications that provide feedback from AI.					
17	I can keep up-to-date with the latest developments in applications that utilize AI for feedback.					
18	I can verify the reliability of AI-generated feedback applications in providing accurate information.					

Article received 2026-02-11