

Development of the Feedback Experience Scale for High School Students

Özlem Koray^{a*}, Seçil Turan^b, Özgür Murat Çolakoğlu^c, & Emine Kahraman^d

a* Prof. Dr., Zonguldak Bulent Ecevit University, <https://orcid.org/0000-0003-1804-0871> ocankoray@gmail.com

b Teacher, Ministry of National Education, <https://orcid.org/0000-0002-0146-5788>

c Asst. Prof. Dr., Zonguldak Bulent Ecevit University, <https://orcid.org/0000-0001-7258-4007>

d Dr., Zonguldak Bulent Ecevit University, <https://orcid.org/0000-0002-0721-9545>

Research Article

Received: 8.10.2021

Revised: 12.9.2023

Accepted: 28.11.2023

Abstract

The quality of feedback in education is important for the development of students. Therefore, the lack of a valid and reliable measurement tool revealing the amount and quality of the feedback students receive during the education process stands out. Within the framework of the study purpose, the Feedback Experience Scale (FES) was developed to fill this gap. The study group consisted of a total of 222 high school students in the first application and a total of 1041 high school students in the second implementation. The data were analyzed by Exploratory Factor Analysis using SPSS and by Confirmatory Factor Analysis using the AMOS program. Item analysis and reliability coefficient were found to perform reliability and validity analyzes of the scale. A two-factor scale with 17 items was developed after the analysis of the scale, which had 23 items and three factors before the implementation. Cronbach Alpha (α) reliability of the scale was .91. Internal consistency coefficients were .89 for the “what you do with the feedback” sub-factor and .63 for the “quantity and timing of feedback” sub-factor. According to the results, the scale is a measurement tool validly and reliably determining students' feedback experience.

Keywords: Feedback, feedback experience, scale development, high school students.

Ortaöğretim Öğrencileri için Geribildirim Deneyimi Ölçeği Geliştirme Çalışması Öz

Eğitimde kullanılan geribildirimlerin kalitesi öğrencilerin gelişimi açısından önemlidir. Bu nedenle, öğrencilerin eğitim sürecinde aldıkları geribildirim miktarı ve niteliğini ortaya koyabilen geçerli ve güvenilir bir ölçme aracının eksikliği dikkat çekmektedir. Çalışmanın amacı kapsamında bu eksikliği gidermek için Geribildirim Deneyimi Ölçeği (GDA) geliştirilmiştir. Araştırmanın çalışma grubunu, ilk uygulama için 222 ortaöğretim öğrencisi; ikinci uygulama için 1041 ortaöğretim öğrencisi oluşturmaktadır. Veriler, SPSS kullanılarak Açıklayıcı Faktör Analizi ve AMOS programı kullanarak Doğrulayıcı Faktör Analizi ile analiz edilmiştir. Ölçeğin güvenirlik ve geçerlik analizleri için madde analizi yapılmış ve güvenirlik katsayısı hesaplanmıştır. Ölçeğin uygulama öncesi üç faktörlü 23 maddeden oluşan madde havuzundan yapılan analizler sonrasında, iki faktörlü 17 madde içeren bir ölçek formu oluşturulmuştur. Ölçeğe ait iç tutarlılık Cronbach alfa katsayısı .91 olarak belirlenmiştir. Analizler sonucunda elde edilen alt faktörler için iç tutarlılık katsayıları; ‘geribildirim’ .89 ve ‘geribildirim miktarı ve zamanlaması’ .63 olarak hesaplanmıştır. Elde edilen sonuçlara göre ölçeğin öğrencilerin geribildirim deneyimlerini geçerli ve güvenilir bir şekilde ortaya koyabilen bir ölçme aracı olduğu söylenebilir.

Anahtar kelimeler: Geribildirim, geribildirim deneyimi, ölçek geliştirme, ortaöğretim öğrencileri.

To cite this article in APA Style:

Koray, Ö., Turan, S., Çolakoğlu, Ö. M. & Kahraman, E. (2024). Development of the feedback experience scale for high school students. *Bartın University Journal of Faculty of Education*, 13(2), 261-273. <https://doi.org/10.14686/buefad.1006391>

INTRODUCTION

Increasing the quality of education or ensuring the maintenance of educational development is an important issue for educators, particularly but for the whole society in general. One of the most important elements that have the capacity to increase the quality of education is measurement and evaluation. Measurement and evaluation are considered as the numerical measurement of the learning outcomes in individuals and the judgment made according to the measurement results. However, if the outcomes individuals gain at the end of the learning process are turned into numerical expressions that are not so meaningful for them, they may not be able to predict the level of their own learning experience. Therefore, the concept of "assessment for learning" has gained importance in recent years instead of the concept of "assessment of learning" (Garcia, Garcia-Alvarez & Moreno, 2014; OECD, 2005).

Assessment is an extremely important process for learning because the individual, who receives feedback about his own learning and is aware of his learning experiences, can use all his energy by focusing on the task given to train himself rather than competing with others to be successful. For this purpose, the main activity of learning is informing students during the learning and teaching process in various ways, using various assessment techniques to check whether the lesson is comprehended or not. Assessment refers to the product assessment in learning environments made by measuring what students learn at the end of a unit and the formative assessment made based on process to determine students' learning needs (Liu & Carless, 2006; OECD, 2005). The formative assessment is an alternative to the type of assessment that students are responsible for not being out of the system in order to continue their future educational life. Formative assessment is based on determining the learning needs of students in the process and teaching them accordingly. According to Shute (2007), in this process, it is necessary to increase students' general skills such as comprehension and problem solving in the areas of knowledge, skill and certain content. Focusing on to what extent how successful the students are in line with the objectives by being responsible for their own learning throughout the teaching process, the main point of formative assessment is providing effective feedback.

If the student is not able to perform the expected performance, the feedback refers to the descriptive information given about the deficiencies and errors the student has in order to close the gap between their current performance (Sonmez, 2007; Tok, 2007). Feedback is used in the education process to determine whether the behaviors students are expected to attain are achieved or not (Hattie & Timperley, 2007). During the feedback process, students try to perform the desired performance by comprehending what is expected of them. This increases students' learning and motivation for the class.

Used to provide information about the learning outcome, corrective feedback helps students to self-assess and thus realize their potential. It also enables students to make self-assessment and feel a sense of accomplishment by realizing their own development (Gibbs & Simpson, 2004). The feedback provided helps students to be aware of their own competencies.

Many studies put forth the effectiveness of the feedback made during and after the learning process (Bergil & Atli, 2012; Black & William, 1998; Espasa & Meneses, 2010; Hu & Choo, 2016; Kleinknecht & Groschner, 2016; Roessger, Daley & Hafez, 2018; Woods & Welch, 2018). In their study, Higgins, Hartley, and Skelton (2002) students tried to learn the feedback that can help them understand the subject deeply by acting with intrinsic motivation even though they were aware of the importance of scoring. The study of Bose and Rengel (2009) revealed that if the students were not given sufficient individual feedback about their own learning, they could not progress because they could not determine at what stage they were in from the point of task-related learning goals, knowledge, and skill. Feedback provided to the students are important because they affect their participation in the class by increasing their motivation and because they help correct their mistakes.

The quality of feedback provided in education is important for the development of students' critical thinking skills or knowledge of something. In this context, the lack of a valid and reliable measurement tool measuring the feedback experiences of students by associating the amount and quality of feedback they received during the education process is remarkable. In order to overcome this shortcoming, this study aimed to develop the Feedback Experience Scale (FES). In the study, it was aimed to reveal the quality and amount of the feedback that the students received during the learning process. This study is important as it is a data source for students' current feedback experiences. Thus, it is believed that the scale will contribute to the teachers, relevant experts and program developers since the results obtained from the scale developed will determine whether or not the students received feedback during the evaluation process, and their views on the quality of the feedbacks if they received any, or their views on their beliefs regarding the effect of feedback on their own learning. In addition, the results

obtained from the scale can be a guide for in-service trainings to be developed for teachers by the Ministry of National Education.

METHOD

In the following section, the methodology of the present study is systematically delineated, aligning with the recognized three stages of scale development, encompassing eight steps as outlined by DeVellis (2012). These stages include design, development, and evaluation, each contributing to the comprehensive description of the study group, the nuanced development process of the FES scale, data collection instruments, and rigorous data analysis techniques. The process adheres to the principles and stages that DeVellis (2012) has detailed, emphasizing conceptual clarity, psychometric validation, and empirical justification. The step-by-step articulation of these phases ensures transparency and replicability, underscoring the study's contribution to the research on feedback experiences within the educational setting.

Design Phase

In the present study, the exploration of assessment and feedback experiences among high school students in grades 9 to 12 led to a focused and robust inquiry. The initial inspiration was drawn from the original Assessment Experience Questionnaire (AEQ) by Gibbs and Simpson (2003), a comprehensive instrument encompassing six distinct dimensions. However, the specific focus of our research led us to concentrate solely on three factors related to feedback: “what you do with the feedback”, “quality of feedback”, and “quantity and timing of feedback”. This decision was guided by the theoretical constructs underpinning our exploration and the aim to investigate the intricacies of feedback within the educational context. In the absence of an existing scale aimed directly at high school students' feedback experiences, researchers examined similar studies and created an item pool accordingly. The design phase commenced with the translation of relevant parts of the AEQ from English to Turkish, carried out by three domain experts. Multiple translated forms were consolidated into a unified version by two additional specialists. The process of creating the items of the FES scale started by initially presenting them to the experts. Based on the suggestions provided by the experts, adjustments were made to each item. Throughout this process, no statistical analysis was applied; instead, the items were validated by the experts before implementation and finalized. This rigorous process culminated in the creation of the Feedback Experience Scale (FES), tailored to resonate with the cultural and educational nuances of the population under study. The FES, focusing on feedback, the quality of feedback, and the amount and timing of feedback, formed the core of our research instrument and was crafted with diligence to align cohesively with the theoretical framework of our study.

Development Phase

In the development phase of this study, a meticulous process was followed to create a 23-item scale tailored to reflect the specific context and needs of the exploration of feedback experiences among high school students. Building upon the foundational work of Gibbs and Simpson's (2003) Assessment Experience Questionnaire (AEQ), the research team identified specific factors crucial for forming the Feedback Experience Scale (FES). A five-point Likert scale was utilized, with responses ranging from “strongly disagree (1)” to “strongly agree (5)”. The initial scale was created in alignment with the relevant literature and the research objectives, and was subsequently evaluated by two assessment and evaluation experts, along with a language expert. Necessary adjustments were made, retaining the 23-item form. Additional items related to the selected feedback factors were thoughtfully incorporated into the item pool, guided by previous research findings and theoretical perspectives. The resulting FES was crafted with diligence, maintaining a strong alignment with the theoretical constructs under investigation, and was administered in schools after obtaining the required permissions. This robust development process ensured that the scale resonated with the unique characteristics of the targeted educational context and population.

Evaluation Phase

The evaluation phase centers on the comprehensive statistical assessment of the AEQ's validity and reliability. In the pursuit of identifying the underlying structure of the scale, a series of factor analyses were conducted, including Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). These analyses provided a strong foundation for the scale's measurement properties, delineating its factor structure in alignment with theoretical expectations. To further validate the instrument, item-total correlations were scrutinized, and Cronbach's alpha internal consistency coefficient was calculated to evaluate reliability. These procedures were meticulously carried out, employing appropriate statistical techniques, to ensure the appropriateness and robustness of the scale in assessing the middle school students' feedback experiences. The

rigorous statistical assessment underpins the scale's applicability and contributes to the broader literature on feedback evaluation within educational contexts.

Study Group

In the study, the 23-item version of the FES was subjected to EFA and CFA procedures to investigate its construct validity. Criterion sampling, one of the purposeful sampling methods, was employed. The study group consisted of high school students attending public schools located in the Western Black Sea Region during the 2018-2019 academic year. In order to ensure the construct validity of the scale, 222 students were included in the study for EFA referred as study group 1 and 1041 students for CFA referred as study group 2. First study group were 222 9th (62), 10th (52), 11th (61) and 12th (47) grade high school students. 112 of the students were female, and 110 were male. In order to examine the consistency between the construct of the scale, which was performed by EFA, and the independent sample, data were collected from a different sample group of 1041 students, of which 470 were female, and 571 were male. Second study group consisted of 9th (269), 10th (280), 11th (250) and 12th (242) grade high school students.

Data Analysis

Before embarking on the detailed process of reliability and validity analysis, preliminary checks and assessments were conducted to ensure the integrity of the data. This began with a thorough examination to confirm the absence of missing data, ensuring a complete and robust dataset for further exploration. Additionally, the evaluation of skewness and kurtosis values was meticulously performed to confirm the normality of the distribution, a critical prerequisite for the application of factor analysis techniques. The EFA was conducted using SPSS version 20, and CFA was carried out with the AMOS statistical package. For all the analyses, the significance level was accepted as .05. An in-depth assessment of the mean and standard deviation of the items provided insights into the central tendency and dispersion of the responses, thus setting the stage for the subsequent stages of reliability and validity assessment. The data analysis was conducted in three primary sections, encompassing reliability, validity, and limitations, to provide a comprehensive evaluation of the Feedback Experience Scale.

Reliability Analysis: Within the scope of the reliability analysis of the Feedback Experience Scale, both item-total correlation and Cronbach Alpha values were examined.

Validity Analysis: The results obtained from two different factor analysis techniques, EFA and CFA, were reviewed for validity analysis.

EFA Technique: Principal axis factoring method was used for extraction of factors of the data, and the “varimax” rotation method was used for factor-item distributions. EFA is a statistical technique that aims to explain the measurement with a few factors by combining variables that measure the same structure or quality (Buyukozturk, 2018; Cecen, 2006). To determine whether the measuring tool is relevant for factor analysis, the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) value and Bartlett’s Test of Sphericity were examined. EFA was conducted with SPSS version 21.

CFA Technique: CFA was conducted with an independent group using AMOS statistical software to verify how well the proposed measurement model from EFA fit the empirical data (Byrne, 2001). The findings include:

- Adequate Fit Indicators: NFI, GFI, AGFI, IFI, and CFI values greater than 0.90 indicate adequate fit.
- RMSEA: Smaller than 0.05 indicates a good fit; below .08 indicates an acceptable fit.
- Chi-square to Degree of Freedom Ratio: Below 5 indicates a good fit, but in this study, the χ^2/df ratio was found to be 8.34, greater than 5. This high value was affected by the large sample size ($\chi^2=983.78$, $N=1041$, $sd=118$, $p=.00$).

Limitations and Considerations:

Although the chi-square value is a basic measurement of the relevance between the model and data, there are limitations in its use. A high χ^2/df ratio may still indicate a good model fit if other fit indices are satisfactory (Kline, 2016). According to Hu and Bentler (1999), χ^2/df can be unreliable, especially in large samples. Therefore, relying on other fit indices like CFI, TLI, and RMSEA is recommended. In large samples ($N>50$), chi-square statistics will be ignored as it mostly rejects relevance between model and data (Bentler & Bonett, 1980; Joreskog & Sorbom, 1993; Kenny & McCoach, 2003).

FINDINGS

In this section, findings related to the validity and reliability works are presented.

Validity of the Feedback Experience Scale

Validity refers to the ability of a test to measure a desired variable without mixing it with another variable (Baykul, 2000; Kane, 2001). Validity in measurement is primarily made to determine whether the scale is appropriate for the characteristics to be measured. Whether the measurement is done in accordance with the rules and whether the measurements reflect the characteristic to be measured are the other issues that require validity in measurement (Mehrens & Lehmann, 1991; Sencan, 2005). In this study, EFA was used to reveal the construct validity for the FES that was intended to be developed. According to Table 1, KMO value of the scale was determined as 0.93. This value shows that the measurements have a sampling adequacy for factor analysis (Sencan, 2005). Bartlett's Test of Sphericity is a hypothesis test to reveal whether the $m \times m$ correlation matrix is an identity matrix. If the $m \times m$ correlation matrix is an identity matrix, it means that the $m \times m$ correlation matrix obtained from the variables is not relevant for factor analysis. According to Table 1, Bartlett's test of sphericity chi-square value was found significant. According to this result, the correlation matrix with $m \times m$ structure is not an identity matrix so that, the data set is adequate for factor analysis (BKTD= 1817.26; $p < 0.05$).

Table 1. KMO and Bartlett's Test Statistics of the Feedback Experience Scale

KMO and Bartlett's Test	Values
Kaiser-Meyer-Olkin (KMO) Sampling Adequacy	.93
Bartlett's Test of Sphericity Approximate Chi-Square Value	1817.26
Degree of Freedom (df)	136
Significance Level (Sig.)	.00

In the factor analysis process, principal axis factoring was performed as a factor extraction method, and also varimax rotation operation was performed in order to simplify the distribution of items under factors. After the factor extraction, there are 17 items left on the scale. The first factor of the FES namely "what you do with the feedback" has 14 items (g1, g2, g3, g4, g5, g6, g7, g8, g9, g10, g11, g12, g13, g14), and the second factor namely "quantity and timing of feedback" has three items (g15, g16, g17). Factor loadings of the items in FES are shown in Table 2.

Table 2. Factor Loadings of the Items in FES

No	Feedback Experience Scale Items	Factors	
		What You Do with Feedback	Quantity and Timing of Feedback
g1	"The feedback makes me aware of the missing information I have on the subject."	.76	
g2	"The feedback shows me how to do better the next time."	.76	
g3	"The feedback makes me aware of my mistakes."	.75	
g4	"The feedback helps me to understand things better."	.74	
g5	"The feedback guides the learning process."	.71	
g6	"The feedback is effective in improving the quality of learning."	.71	
g7	"The feedback helps me in improving whatever I need to improve in a task (assignment)."	.68	.33
g8	"Once I have read the feedback I understand why I got the mark I did."	.60	
g9	"The feedback essentially helps me understand how well I did compared to others."	.59	
g10	"I carefully read the feedback and try to understand what it says."	.59	
g11	"I use feedback while correcting my assignment."	.57	
g12	"Getting feedback after exam makes me learn."	.53	
g13	"The feedback prompts me to go back over material covered earlier in the course."	.48	
g14	"The feedback helps me with the subsequent assignments."	.43	
g15	"I get feedback in many courses each semester."		.72
g16	"In many courses, I receive feedback on my level of success."		.68
g17	"The feedback comes very quickly."		.64

As seen in Table 2, factor loadings of “what you do with feedback” ranged between .43 to .76. Similarly, factor loadings of “quantity and timing of feedback” ranged between .64 to .72.

According to Table 3, the reliability coefficient of the two sub-factors of the FES is 0.89 and 0.63 respectively. The first factor of the measurement tool was named as “what you do with feedback”, and the eigenvalue of this factor is 7.65. The percentage of variance explained for the first factor was found 44.98%. The second factor was named as “quantity and timing of feedback”, and the eigenvalue of this factor is 1.48. The percentage of variance explained for the first factor was found as 8.72%.

Table 3. Reliability Coefficients of the Feedback Experience Scale Sub-factors

	What You Do With Feedback (1 st Factor)	Quantity and Timing of Feedback (2 nd Factor)
Cronbach Alfa (α)	0.89	0.63
Eigenvalues of Factors	7.65	1.48
Explained Variance Percentage	44.98	8.72
Explained Cumulative Variance Percentage	44.98	53.69

It is seen that the two-factor structure of the FES explains 53.69% of the feedback experience. When the psychometric properties of the scale are examined, it can be said that the scale is a valid and reliable measurement tool for determining the feedback experiences of students at high school.

Confirmatory Factor Analysis

After the EFA process, the last version of the scale was conducted to CFA on a different sample group. For the validity of CFA results, fit indices of the model must meet the required cutoff criteria's.

The RMSEA value between 0 and 0.05 suggests a good fit value, and value between 0.05 and 0.08 means an acceptable fit (Brown & Cudeck, 1993; Byrne & Campbell, 1999). The RMSEA value of the CFA model is found to be 0.08, which means an acceptable fit between model and data. The other fit statistics calculated in this analysis such as CFI, NFI and IFI were found respectively 0.86, 0.84 and 0.85, although these values were close to .90, they were all below this value. According to the calculations, the GFI value being higher than 0.85 and the AGFI value being higher than 0.80 can be accepted as justification for the fit of the model with the data (Anderson & Gerbing, 1984; Cole, 1987; Gulbahar & Buyukozturk, 2018; Marsh, Balla & McDonald, 1988). For this study, the values of GFI and AGFI were found to be 0.89 and 0.85, respectively. The relevant fit indices CFI, GFI, NFI, and TLI are important indicators commonly used to evaluate model fit. However, if all of these indices fall below .9, the model may be considered to exhibit an unacceptable fit. In such cases, researchers may work on alternatives such as restructuring the model or considering some variables in the data set. However, in many cases, it may not be possible for the values of these fit indices to be exactly .9 or above. In a study conducted by Hu and Bentler (1999), it was shown that the model fit could still be acceptable even if the values of these indices were below .90. Therefore, it may be more appropriate to consider other factors in addition to the fit indices when making a decision rather than simply rejecting a model based solely on fit indices.

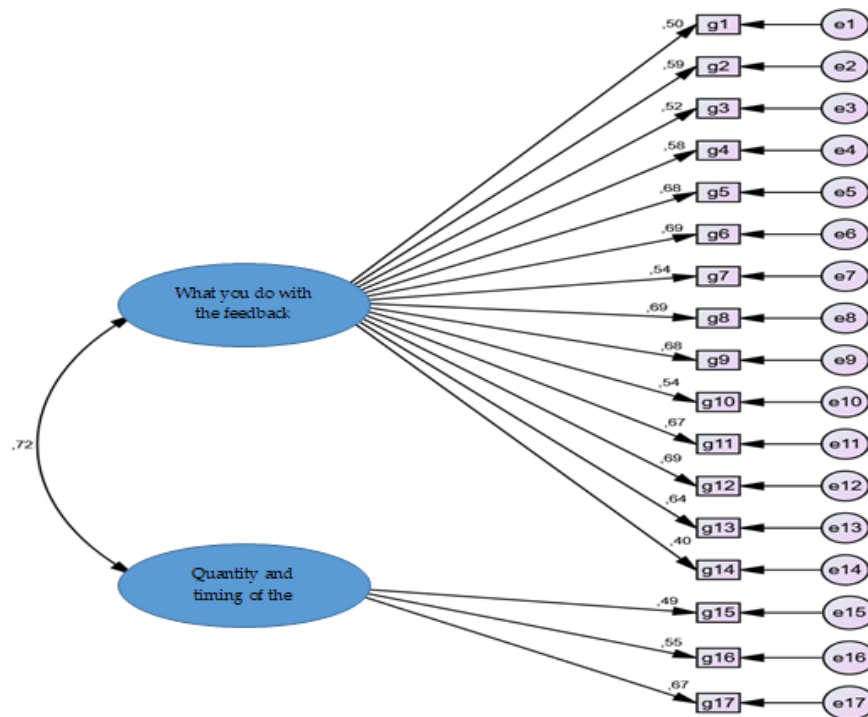


Figure 1. *Confirmatory Factor Analysis Model*

In the Confirmatory Factor Analysis conducted to validate the Feedback Experience Scale, the resulting diagram, depicted in Figure 1, confirms a two-dimensional structure of the scale, with factor loadings ranging between 0.40 and 0.69. The final version of the scale demonstrates acceptable fit indices, and notably, all factor loadings were found to be statistically significant at the 0.05 level.

Reliability Scores of Feedback Experience Scale

After the validity studies of a scale, reliability analyses should be conducted. In this section, first of all, the descriptive findings of the feedback experience scale for secondary school students were examined. The findings regarding the mean, standard deviation and item-total correlation values of the items are shown in Table 4. Additionally, the item-total correlation coefficients of all items are given in Table 4. Item-total correlation is given because it explains the relationship between the scores obtained from the test items and the total score of the test. Positive and high item-total correlation indicates that the items sample similar characteristics. In general, it can be said that items with an item-total correlation below .20 should not be included in the test (Buyukozturk, 2018).

Table 4. Descriptive Statistics for the Items in the Feedback Experience Scale

No	Feedback Experience Scale Items	Mean	Standard deviation	Item-total Correlation
g10	“I carefully read the feedback and try to understand what it says.”	3.84	.98	.59
g14	“The feedback helps me with the subsequent assignments.”	3.72	2.26	.42
g13	“The feedback prompts me to go back over material covered earlier in the course.”	3.38	1.09	.54
g11	“I use feedback while correcting my assignment.”	3.72	1.08	.60
g1	“The feedback makes me aware of the missing information I have on the subject.”	3.91	1.03	.69
g5	“The feedback guides the learning process.”	3.78	1.06	.68
g9	“The feedback essentially helps me understand how well I did compared to others.”	3.47	1.06	.59
g4	“The feedback helps me to understand things better.”	3.76	.97	.71
g2	“The feedback shows me how to do better the next time.”	3.75	1.04	.74
g8	“Once I have read the feedback I understand why I got the mark I did.”	3.78	1.11	.62

g6	“The feedback is effective in improving the quality of learning.”	3.63	1.06	.73
g3	“The feedback makes me aware of my mistakes.”	3.92	1.04	.70
g7	“The feedback helps me in improving whatever I need to improve in a task (assignment).”	3.61	1.13	.73
g16	“In many courses, I receive feedback on my level of success.”	3.00	1.19	.49
g17	“The feedback comes very quickly.”	2.82	1.08	.43
g15	“I get feedback in many courses each semester.”	3.04	1.20	.45
g12	“Getting feedback after exam makes me learn.”	3.69	1.16	.57

According to Table 4, the item-total correlation for 17 items varies between .42 and .74. As can be seen, all of the values are exceeded the recommended cut off point of 0.3 (Field, 2013). This result indicates that the Feedback Experience Scale is associated with the measured properties and thus enhances the reliability of the scale. Additionally, it demonstrates that all items were selected appropriately for the purpose of the scale and reflect the measured properties.

The examination of the means and standard deviations of the items in the Feedback Experience Scale reflects students' overall thoughts on various feedback experiences. Looking at the means of the items, it can be seen that most statements are within the range of 3-point values. This indicates that students tend to think that feedback is generally useful and plays an important role in the learning process.

Regarding the standard deviations, values slightly above “1” are observed for most items. This shows that students' responses to these statements exhibit a more central tendency, meaning that most students give similar responses to these items. However, the standard deviation of 2.26 for the item “The feedback helps me with the subsequent assignments” suggests that responses to this item have a wider distribution. This may indicate that students' thoughts on this issue may be more diverse.

DISCUSSION & CONCLUSION

The primary aim of this research was to devise a reliable and valid scale that measures and evaluates the feedback experiences of Turkish high school students in grades 9 to 12. It specifically emphasizes the quality, quantity, timing, and interaction of feedback a fundamental aspect of student learning that enhances understanding, challenges overcoming, and attainment of learning goals (Black & Wiliam, 1998; Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Shute, 2008). Despite the prominence of feedback in education, the literature reveals a conspicuous gap in assessing both the effects of the feedback process on student achievement and the students' experiences in this process, especially in Turkey. This study endeavors to address this lacuna by meticulously crafting a valid instrument for teachers, administrators, experts, and policymakers, thereby enriching our grasp on student needs and strategizing to render the feedback process more effective and purpose-driven.

Feedback's role in education extends beyond mere evaluation and can profoundly impact learning permanence (Hattie & Timperley, 2007). Various studies have explored feedback's nature, including examinations of teacher comments (Brookhart & Moss, 2009; Chase & Houtmanfa, 2009; Mutch, 2003) and comparisons between basic and detailed feedback (Chase & Humanas, 2009). Literature emphasizes timely, specific, accurate, and purposeful feedback that contributes positively to student performance (Burnett & Mandel, 2010; Irons, 2008; Lee, 2017; Noor, Aman, Mustaffa & Seong, 2010; Poulos & Mahony, 2008). Indeed, feedback's influence reaches beyond content, affecting students' overall learning experience (Askew & Lodge, 2000; Butler & Winne, 1995; Guven, 2004; Higgins, Hartley, & Skelton, 2002; Ilgen, Fisher & Taylor, 1979; Lipnevich & Smith, 2008; Nicol & Macfarlane Dick, 2006; Panasuk & Lebaron, 1999; Peterson & Irving, 2008; Poulos & Mahony, 2008). In this broad literary context, the current study aims to develop a tool for measuring and evaluating the multidimensional aspects of feedback experiences.

Also, the study conducted by Akkuzu & Uyulgan (2014) focused on developing a feedback scale to measure prospective teachers' thoughts and affective reactions. Two factors, namely professional development and anxiety, highlight the multifaceted nature of feedback. The emphasis on professional development resonates with the current study's focus on enhancing student learning, while the exploration of anxiety adds a unique dimension, reflecting specific concerns for prospective teachers.

The focus on feedback in this study presents an interesting parallel with the dimensions "What You Do with Feedback" and "Quantity and Timing of Feedback" identified in our research. Akkuzu & Uyulgan's (2014) dimension of professional development may articulate the complexity of how feedback affects learning and

personal growth, and how it aligns with the ability to overcome challenges. Likewise, the anxiety dimension can aid in understanding students' emotional reactions to the timing and amount of feedback.

Additionally, another research by Onlu, Abdusselam & Yilmaz (2022) identified three dimensions: mastery, positive affect, and negative affect. The emphasis on mastery aligns with our focus on feedback as a means of overcoming challenges and attaining learning goals. The recognition of positive and negative affect underscores the affective reactions and the overall learning experience that are congruent with our exploration of the "Quantity and Timing of Feedback" dimension.

In this rigorous academic endeavor, the Feedback Experience Scale (FES) was subjected to validity and reliability analyses, correlating the multi-dimensional aspects of feedback experienced during the education process. Six items were excluded, resulting in a robust 17-item scale with two factors, utilizing a 5-point Likert scale. A higher total score on the scale signifies a more positive feedback experience, aligning with the principles of timely, specific, accurate, and purposeful feedback. The Cronbach Alpha (α) reliability was found to be .91, while the Exploratory Factor Analysis (EFA) affirmed the construct validity, unveiling the structure encapsulating students' feedback experiences.

This scale stands as a novel contribution to the existing body of knowledge, tailor-made for the Turkish context. Educators can leverage the FES to pinpoint areas requiring refinement. For instance, a deficiency in feedback details could prompt a more comprehensive approach. Administratively, the scale could inform policies or training programs, possibly culminating in standardized feedback guidelines or teacher development modules across various educational institutions. Additionally, the FES might find applicability across diverse learning landscapes, such as online education or vocational training. Hypothetical case studies could elucidate how an educational district utilized the scale to overhaul feedback practices, delineating resultant enhancements. Consideration of the cultural and social milieu is paramount, as feedback mechanisms often mirror regional societal norms, values, and traditions (Boud & Molloy, 2013). Specialized studies exploring these interconnections are encouraged, offering fertile ground for further research.

Moreover, employing methodologies such as Item Response Theory (IRT) and cross-validation with other psychometric tools could enrich the instrument's reliability and validity, creating a more vibrant understanding of feedback. This methodological triangulation aligns theoretical underpinnings with pragmatic applications, fostering a more adaptive feedback environment for students across varied learning contexts.

In conclusion, the developed scale stands as an innovative tool to gauge students' feedback experiences, taking into account an array of demographic variables. This study sets a foundational precedent for future research, illuminating feedback's role across educational stages and contributing significantly to our comprehension of feedback's vital function in education. The meticulous work conducted herein promises to facilitate educational stakeholders in assessing and enhancing feedback quality, thereby propelling the field toward a more informed and student-centered approach.

Statements of Publication Ethics

The ethics committee report of this research was obtained from Bulent Ecevit University Human Research Ethics Committee (Date: 29/03/2019, Protocol number: 540).

Researchers' Contribution Rate

All authors have participated sufficiently in the work to take public responsibility for the content, research design, analysis, methodology, data collection, resources, discussion, conclusion, writing - review & editing.

Conflict of Interest

There are no conflicts of interest in this study.

REFERENCES

- 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. DOI: 10.1080/03323315.2014.923184.
- Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49, 155-173. <https://link.springer.com/content/pdf/10.1007/BF02294170.pdf>
- Askew, S., & Lodge, C. (2000). Gifts, ping-pong and loops-linking feedback and learning. In S. Askew (Eds.), *Feedback for learning* (pp.1-17). Routledge.
- Baykul, Y. (2000). *Egitimde ve psikolojide olcme: Klasik test teorisi ve uygulamasi*. OSYM Publishing.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588.
- Bergil, A. S., & Atli, I. (2012). Different perspectives about feedback on teaching. *Procedia-Social and Behavioral Sciences*, 46, 5833-5839. doi: 10.1016/j.sbspro.2012.06.524
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5 (1). http://wiki.biologyscholars.org/@api/deki/files/1278/%3Dblack_p_et al_1998b%5B1%5D.pdf
- Bose, J., & Rengel, Z. (2009). A model formative assessment strategy to promote student-centered self-regulated learning in higher education. *US-China Education Review*. 6(12), 29-35. <https://files.eric.ed.gov/fulltext/ED511170.pdf>
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698-712.
- Brookhart, S. M., & Moss, C. M. (2009). *Advancing formative assessment in every classroom: A guide for instructional leaders*. ASCD.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage Focus Editions*, 154, 136-136.
- Burnett, P. C., & Mandel, V. (2010). Praise and feedback in the primary classroom: Teachers' and students' perspectives. *Australian Journal of Educational & Developmental Psychology*, 10, 145-154.
- Butler, D., & Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245-281.
- Buyukozturk, S. (2018). *Sosyal bilimler icin veri analizi el kitabi*. Pegem A Publishing.
- Byrne, B. M. (2001). Structural equation modeling with AMOS, EQS, and LISREL: Comparative approaches to testing for the factorial validity of a measuring instrument. *International journal of testing*, 1(1), 55-86. doi: 10.1207/S15327574IJT0101_4,
- Byrne, B. M., & Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure: A look beneath the surface. *Journal of Cross-Cultural Psychology*, 30(5), 555-574. <https://doi.org/10.1177/0022022199030005001>
- Cecen, A. R. (2006). Duygulari yonetme becerileri olceginin gelistirilmesi: Gecerlik ve guvenirlik calismalari. *Turkish Psychological Counseling and Guidance Journal (TPCGJ)*, 3(26), 101-113.
- Chase, J. A., & Houmanfar, R. (2009). The differential effects of elaborate feedback and basic feedback on student performance in a modified, personalized system of instruction course. *Journal of Behavioral Education*, 18(3), 245-265. doi 10.1007/s10864-009-9089-2
- Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, 55(4), 1019-1031. <https://doi.org/10.1037/0022-006X.55.4.584>
- Espasa, A., & Meneses, J. (2010). Analyzing feedback processes in an online teaching and learning environment: an exploratory study. *Higher education*, 59(3), 277-292. doi: 10.1007/s10734-009-9247-4

- Devellis, R. F. (2012). *Scale development theory and practice* (3rd ed.). Sage Publications, Inc.
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. Sage Publications.
- García, A. S., García-Álvarez, M. T., & Moreno, B. (2014). Analysis of assessment opportunities of learning spaces: On-line versus face to face methodologies. *Computers in Human Behavior*, 30, 372-377. <https://doi.org/10.1016/j.chb.2013.04.024>
- Gibbs, G., & Simpson, C. (2003). Measuring the response of students to assessment: The Assessment Experience Questionnaire. *11th Improving Student Learning Symposium*, 1-12. <https://pdfs.semanticscholar.org/5711/1701f9d713bde7780f6442e6bb4c1cf1b43b.pdf>
- Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3-31. <http://eprints.glos.ac.uk/3609/1/LATHE%201.%20Conditions%20Under%20Which%20Assessment%20Supports%20Stud>
- Gulbahar, Y., & Buyukozturk, S. (2008). Adaptation of assessment preferences inventory to Turkish. *Hacettepe University Journal of Education*, 35, 148-161.
- Güven, İ. (2004). Etkili bir öğretim için öğretmenlerden beklenenler. *Milli Eğitim Dergisi*, 164, 127– 141.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81–112. doi: 10.3102/003465430298487
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27(1), 53-64. doi:10.1080/03075070120099368
- Hu, G., & Choo, L. (2016). The impact of disciplinary background and teaching experience on the use of evaluative language in teacher feedback. *Teachers and Teaching*, 22(3), 329-349. doi:10.1080/13540602.2015.1058591
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>.
- Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), 349-371.
- Irons, A. (2008). *Enhancing learning through formative assessment and feedback*. Routledge.
- Joreskog, K. G., & Sorbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Scientific Software International.
- Kane, M.T. (2001), Current Concerns in Validity Theory. *Journal of Educational Measurement*, 38: 319-342. <https://doi.org/10.1111/j.1745-3984.2001.tb01130.x>
- Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling*, 10(3), 333-351. doi: 10.1207/S15328007SEM1003_1
- Kleinknecht, M., & Groschner, A. (2016). Fostering preservice teachers' noticing with structured video feedback: Results of an online-and video-based intervention study. *Teaching and Teacher Education*, 59, 45-56. <http://dx.doi.org/10.1016/j.tate.2016.05.020>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer.
- Lipnevich, A. A., & Smith, J. K. (2008). I really need feedback to learn: Students' perspectives on the effectiveness of the differential feedback messages educational assessment. *Evaluation and Accountability*, 21(4), 347-367. doi: 10.1007/s11092-009-9082-2
- Liu, N. F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290. doi: 10.1080/13562510600680582

- Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391-410. <https://files.eric.ed.gov/fulltext/ED267091.pdf>
- Mehrens, W. A., & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology*. Forth Worth.
- Mutch, A. (2003). Exploring the practice of feedback to student. *Active Learning in Higher Education*, 4(1), 24-38.
- Nicol, D. J., & Macfarlane Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. doi: 10.1080/0307570600572090
- Noor, N. M., Aman, I., Mustaffa, R., & Seong, T. K. (2010). Teacher's verbal feedback on students' response: a Malaysian ESL classroom discourse analysis. *Procedia-Social and Behavioral Sciences*, 7(0), 398-405. doi:10.1016/j.sbspro.2010.10.054
- OECD, (2005). Formative assessment: Improving learning in secondary classrooms. *Policy Brief*, 1-8. www.oecd.org/publications/Policybriefs.
- Onlu, O. B., Abdusselam, M. S., & Yilmaz, R. M. (2022). Students' Perception of Instructional Feedback Scale: Validity and Reliability Study. *Contemporary Educational Technology*, 14(3), ep368.
- Panasuk, M. R., & Lebaron, J. (1999). Student feedback: A tool for improving instruction in graduate education. *Education*, 120(2), 356-368.
- Peterson, E. R., & Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction*, 18(3), 238-250. <https://doi.org/10.1016/j.learninstruc.2007.05.001>
- Poulos, A., & Mahony, M. J. (2008). Effectiveness of feedback: The students' perspective. *Assessment & Evaluation in Higher Education*, 33(2), 143-154. <https://doi.org/10.1080/02602930601127869>
- Roessger, K. M., Daley, B. J., & Hafez, D. A. (2018). Effects of teaching concept mapping using practice, feedback, and relational framing. *Learning and Instruction*, 54, 11-21. <https://doi.org/10.1016/j.learninstruc.2018.01.011>
- Sencan, H. (2005). *Sosyal ve davranissal olcumlerde guvenirlik ve gecerlilik*. Seckin Publishing.
- Shute, V. J. (2007). *Focus on formative feedback*. Research Report. Educational Testing Service. <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2333-8504.2007.tb02053.x>
- Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), 153-189. <https://doi.org/10.3102/0034654307313795>
- Sonmez, V. (2007). *Program gelistirmede ogretmen el kitabı*. Ani Yayıncılık.
- Tok, S. (2007). *Ogretme-ogrenme strateji ve modelleri, ogretim ilke ve yontemleri*. Pegem A Publishing.
- Woods, A., & Welch, T. D. (2018). Feedback board: an innovative way to adapt classroom teaching. *Teaching and Learning in Nursing*, 13(4), 219-222. <https://doi.org/10.1016/j.teln.2018.05.006>

APPENDIX**Appendix A. Items of the Turkish Form of Feedback Experience Scale**

No	Feedback Experience Scale Items
1	Geribildirimler konu hakkında eksik bilgilerimin farkına varmamı sağlar.
2	Geribildirimler çalışmamın bir dahaki sefere nasıl daha iyi yapabileceđimi gösterir.
3	Geribildirimler hatalarımın farkına varmamı sağlar.
4	Geribildirimler bana bir şeyleri daha iyi anlamam konusunda yardım eder.
5	Geribildirimler öğrenme sürecinde yol göstericidir.
6	Geribildirimler öğrenmenin kalitesinin artırılmasında etkilidir.
7	Bir çalışmada (ödevde) neyi geliştirmem konusunda geribildirimler bana yardımcı olur.
8	Geribildirimi okuduđum zaman neden o puanı aldıđımı anlarım.
9	Geribildirimler, temelde diđerlerine göre ne kadar iyi yaptıđımı anlamamı sağlar.
10	Geribildirimi dikkatlice okur ve ne söylediđini anlamaya çalışırım
11	Ödevimde düzeltme yaparken geribildirimleri kullanırım.
12	Sınav sonrasına geribildirim almam öğrenmemi sağlar.
13	Geribildirim, derste kullanmış olduđum materyallere tekrar göz atmam için beni yönlendirir.
14	Geribildirimler sonradan verilen ödevlerde bana yardımcı olur.
15	Her dönem birçok derste geribildirim alırım.
16	Birçok derste, başarı düzeyim konusunda geribildirim alırım.
17	Geribildirimlerin bana dönüşü oldukça hızlıdır.
