Journal of Education and Future
year: 2017, issue: 12, 117-135

Development and Validation of Formative Feedback Perceptions Scale in Project Courses for Undergraduate Students

Mustafa Şat*

Abstract

There has been an increasing trend towards project-based approach in higher education. For that reason, assessing and evaluating students' perceptions on formative feedback becomes an important aspect of students' learning in groupwork. Valid and reliable instruments are required to measure how students perceive formative feedback given to them from multiple sources. Therefore, the aim of the current study is to develop and validate a multidimensional formative feedback perception scale (FFPS) for the assessment of undergraduate students' perceptions on formative feedback provided to them as individuals or as a group on their project work. This study was implemented in three phases. In the first phase, qualitative data through interviews with 10 participants were used in development of the items for the scale. In the second phase, quantitative data from 97 undergraduate students were used to ensure validity and reliability of the new scale through exploratory factor analysis. Three dimensions emerged and were interpreted as development, understandability, and encouragement. In the third phase, data collected from an independent sample of 250 undergraduate students were used to confirm the factorial structure of the 25-item FFPS through confirmatory factor analysis. This new scale will serve as a useful tool for researchers, practitioners andindividuals in higher education in order to assess students' formative feedback perceptions in various work-based and experiential learning environments.

Keywords: formative feedback perceptions, scale development, undergraduate students, projects-based course

* PhD Student, Middle East Technical University, Faculty of Education, Computer Education and Instructional Technology, Ankara, Turkey. E-mail: mstfsat@gmail.com

Lisans Öğrencileri İçin Proje Derslerinde Verilen Biçimlendirici Geri Bildirim Algı Ölçeğinin Geliştirilmesi ve Geçerliği

Öz

Yükseköğretim ortamlarında proje temelli yaklaşıma yönelik gittikçe artan bir eğilim oluşmuştur. Bu nedenle, öğrencilerin biçimlendirici geri bildirim üzerine algılarını değerlendirmek ve ölçmek, öğrencilerin grup çalışmasında öğrenmelerinin önemli bir parçası haline gelmiştir. Öğrencilerin çoklu kaynaklardan kendilerine verilen biçimlendirici geri bildirimi nasıl algıladıklarını ölçmek için geçerli ve güvenilir bir ölçek geliştirme gereği oluşmuştur. Önemli olan bu ihtiyacı karşılamak için, bu çalışmanın amacı, lisans öğrencilerinin proje çalışması sırasında bireysel veya grup olarak aldıkları biçimlendirici geri bildirim hakkındaki algılarına yönelik çok boyutlu bir ölçek geliştirmektir. Üç aşamadan oluşan bu çalışmanın ilk aşamasında 10 katılımcı ile görüşülmüş, elde edilen nitel veriler kullanılarak ölçek maddeleri geliştirilmiştir. İkinci aşamada ölçeğin geçerliği ve güvenirliği sağlamak için 97 lisans öğrencisinden elde edilen nicel veriler açımlayıcı faktör analizi ile incelenmiştir. İkinci aşamada ölçeğin geçerliği ve güvenirliğini sağlamak için 97 lisans öğrencisinden elde edilen nicel veriler açımlayıcı faktör analizi ile incelenmiştir. Analiz sonunda ortaya çıkan üç boyut "gelişme", "anlaşılırlık ve uygunluk" ve "teşvik etme" olarak isimlendirilmiştir. En son aşamada ise 250 öğrenciden oluşan farklı bir desenden elde edilen veriler kullanılarak 25 maddelik ölçeğin faktör yapısı doğrulayıcı faktör analizi ile incelenmiştir. Geliştirilen bu ölçeğin grup çalışmasına dayalı çeşitli öğrenme ortamlarında öğrencilerin biçimlendirici geri bildirim algılarını ölçmek için araştırmacılara, uygulayıcılara ve akademisyenlere yarar sağlayacağı düşünülmektedir.

Anahtar Sözcükler: biçimlendirici geri bildirim algısı, ölçek geliştirme, lisans öğrencileri, proje tabanlı ders

Introduction

Project-based learning (PBL) is a student-centered pedagogy basically formed on the concept of "learning by doing" which was outlined by Dewey in 1897. It has been used pervasively across disciplines as a 21st-century strategy for helping learners actively explore real-world problems and challenges through investigating and collaborating. PBL is described as a learning strategy design that organises learning around the project (Thomas, 2000). An instance of PBL can be characterised by distinctive features such as being central to curriculum, including problems that link activities with underlying conceptual knowledge, involving enquiry process of constructing and transforming knowledge, consisting of learning settings which are mainly driven by students, and incorporating authentic and realistic challenges (Thomas, 2000; Helle, Tynjälä, & Olkinuora, 2006). PBL environment is based on a constructivist approach, designed to promote student-centred learning where the instructor and teaching assistants act as facilitators and mentors, and students are encouraged to deal with problems encountered during learning process. Students are motivated to apply acquired knowledge and skills to achieve learning objectives through constructing their own knowledge and understanding, working independently, or by collaborating in teams, through interaction with an instructor, teaching assistants, and their peers (Thomas, 2000). In accomplishing those objectives in PBL, continuous formative assessment, revision, and evaluation of student performance on project-related teams are quite important.

Previous studies considered formative feedback as a prominent means to advance learning (Hattie, 2008; Wiliam, 2011) and suggested that the role of formative feedback in PBL is significant (Fernandes, Flores, & Lima, 2012; Sadler, 1998). In addition, comprehensive literature reviews on formative assessment revealed that formative feedback facilitates and enhances self-regulated learning (Nicol & Macfarlane- Dick, 2006; Sadler, 1989) by helping students deeply involved in meta-cognitive strategies such as personal goal-planning, monitoring, and reflection (Clark, 2012), and therefore plays a crucial role in the assessment process at the centre of formative assessment (Black & Wiliam, 1998). This makes feedback both central to student learning and also very powerful in facilitating learning (Bandura, 1991).

Certain types of qualitative and quantitative human assessment and judgement are in play when any type of feedback is given on students' team works by their instructor or teaching assistant(s). Feedback (i.e. summative and formative) is asserted to be a process of communication (Higgins, Hartley, & Skelton, 2001) in which the quality of student performance related to the work is evaluated (Sadler, 1989). However, not all feedback can necessarily be considered as formative in nature. Formative feedback described in this study refers to evaluative or corrective information given to individual members of a group or to the group as a whole (London & Sessa, 2006). Taras (2002) puts forward some characteristics in order for feedback to be considered formative. For example, understanding of students (as individuals or as a group) regarding assessment and judgement of their work and how they can avoid deficiencies in the future. Moreover, a subsequent piece of work

that is required to be produced by the students with the problems successfully addressed and remedied.

The nature of the task, the students' needs, and learning environment can result in feedback with different functions (Knight & Yorke, 2003; Poulos & Mahony, 2008). For some, feedback is information aiming to close the gap between current performance and expected performance (Ramaprasad, 1983; Sadler, 1989; Lizzio & Wilson, 2008; Draper, 2009; Wiliam, 2011) by reforming the student's understanding so that learning improves (Shute, 2008). For others, in addition to nurturing ongoing performance towards a target standard, feedback functions to move the current work forwards and to influence the quality of subsequent work (Schraw, Crippen, & Hartley, 2006; Hounsell et al., 2008; Boud & Molloy, 2013). In her seminal article, Ramaprasad (1983, p. 4) defined feedback as "information about the gap between the actual level and the reference level of a system parameter which is used to alter the gap in some way". The conception of formative feedback held by this study corresponds to the aforementioned notions and relies heavily on the definition of Ramaprasad (1983), just as it is convenient to refer to "formative feedback" as noted by Taras (2005).

Several studies have attempted to develop conventional and standardised feedback perception scales in order to elaborate on the feedback issue in conventional classroom settings (Lizzio & Wilson, 2008; Rowe & Wood, 2008). However, conventional and standardised feedback perception scales, where feedback is measured as an assessment of individual student performance, may not be appropriate to fully understand formative feedback as provided on team project works. This is because, in PBL, strengths and weaknesses from the students' point of view involves assessment of the multimode of team performance, skills, and ideas with formative feedback given as multidimensional forms from multiple sources. Therefore, the purpose of this study is to develop a multidimensional scale that can be used to measure undergraduate students' formative feedback perceptions not only in PBL settings within science programmes but also in engineering programmes.

Context of the Study

Computer Education and Instructional Technology undergraduate programme offers several project-based courses with very strong emphasis on project design and development. These courses are designed based constructivist approach where students construct their own knowledge by interacting with staff and colleagues. In general, there are two types of project carried out by students throughout the semester: small-scale and large-scale. While small-scale projects span two to four weeks, large-scale projects span from five weeks to two semesters. Students commonly utilize various cutting-edge educational software and programs as well as well-known instructional design models (e.g., ASSURE, ADDIE model) in order to design and develop their educational product. The completion of the course is depended on development of an education project with acquired knowledge and skills during the course, comprehensive written tasks such as a project report and portfolio and a multimedia presentation. While working on the project, students, as a

group or individuals, meet with staff (instructor and teaching assistants) regularly each week. During these meetings they discuss the flow of their projects, written reports and any issues they faced during the process and receive both written and oral comments regarding their progress. This face-to-face interaction and dialogue also continues in electronic form within online environments through e-mail, open courseware, and social media, especially Facebook. The primary forms of feedback provided during project work sessions was oral feedback given in face-to-face meetings, and written feedback given on draft project work and reports. The roles of the instructor and teaching assistants are to coordinate teams, support them in solving their problems and to monitor the development of the project by providing students with oral and written formative feedback with regards to their achievement and performance.

Method

Samples

Two independent samples of senior and junior undergraduate students from the same programme were recruited from four Turkish public universities in order to provide psychometric evidence for the formative feedback perceptions scale (FFPS). FFPS was developed into two phases: a pilot and the main study. While Sample 1 was used in the pilot study in order to identify and reveal factor structure for the FFPS, Sample 2 was used in the main study to perform cross-validation of the factor structure of FFPS. Sample 1 consisted of 53 senior and 44 junior undergraduates (N = 97; 63 males and 34 females), whilst Sample 2 consisted of 112 senior and 138 junior undergraduates (N = 250; 109 males and 141 females) (see table 1). Convenience sampling method was used for the selection of both Sample 1 and Sample 2. The questionnaire was distributed during course hours with the help of lecturers at the universities. The procedures of the current study were approved by Middle East Technical University Human Subject Ethics Committee.

Table 1

Descriptive Statistics of Participants

	Class		Gender		
	3rd	4th	Male	Female	Total
Sample 1					
Middle East Technical University	44	53	63	34	97
Sample 2					
Amasya University	36	29	23	42	65
Ankara University	47	38	40	45	85
Hacettepe University	55	45	46	54	100

Instrumentation

Participant selection for semi-structured interviews

In this study, the FFPS was designed to examine undergraduate students' perceptions exclusively on formative feedback given to them for their projects and project-related documents. Therefore, purposive sampling was adopted to select

participants for the semi-structured interviews. The criterion for sampling was that the participants should be among the most successful 20% of the students who had recently taken the "Instructional Feedback Design and Development" elective course, because in-depth understanding of the concept of "formative feedback" was deemed necessary for the interviews. Data were collected from the university that offered this course. Among 15 students identified by the instructor based on the criterion, 10 students agreed to participate (seven male, three female).

Formative feedback perceptions questionnaire

The instrument development process took place in several phases. In the first phase, contingent decisions were made based upon the constructs underlying formative feedback perceptions. For this purpose, a comprehensive in-depth relevant literature review was conducted and the previously described constructs underlying feedback perceptions in the context of higher education were identified and noted. During iterative review, particular attention was paid to items in previously developed feedback-related questionnaires due to being convenient for this study. In the second phase, the qualitative dataset collected through the interviews were examined in detail, based on each theme and corresponding subtheme, in order to detect and select, and then each questionnaire item was formulated and written. Qualitative data analysis culminated in the exploration of main themes which were then used to merge with the dimensions of feedback perceptions identified as a result of the literature review. Based on the decisions after merging, three dimensions that addressed different aspects of formative feedback perceptions were proposed: development, encouragement, and understandability. After the three dimensions had been decided, an initial item pool composed of 40 items was generated, each representing different dimensions underlying the formative feedback perception. In the third phase, experts and specialists in instructional feedback practices examined the item pool of FFPS related to content and face validity. Independent cognitive interviews were conducted with one graduate student and two undergraduate students in order to pretest the FFPS. Based on the feedback from the cognitive interviews, except for a few items which were reworded to ensure better readability and understandability of the items, those remaining were accepted without revision. The scale items were measured on a 5-point Likert-type rating scale, ranging from "never" through to "very frequently". Demographic information including gender, university, education level, and cumulative average was also captured.

Specifically, in writing the items, interview data analysis results were predominantly used and supported by relevant studies identified from the literature. While several items were taken directly from interview transcripts, some items were adapted and rewritten by the researcher according to the guidelines of Groves et al. (2009), Dillman (2007), and Oppenheim (1992).

Since the medium of instruction at one university is English and the other three is Turkish, the questionnaire was developed both in the English language (see Appendix) and also in Turkish. For content validity, both versions of the questionnaire were reviewed for their equality by an ESL English Language

instructor. Furthermore, one instructor who has considerable expertise in the subject of feedback practices contributed heavily to every step of instrument development such as final decisions on writing and selection of the items. In addition, two independent instructors with expertise in questionnaire design and the content of the questionnaire also reviewed the questionnaire items for content validity of the instrument. Based on the feedback from the experts, several items were re-worded and some items refined to ensure they were clear and relevant to the formative feedback context.

Data Analysis

The data analysis process was accomplished in four steps. Specifically, it started with identification of the factor structure of FFPS through the use of exploratory factor analysis (EFA), and continued with cross-validation of the factor structure with confirmatory factor analysis (CFA), and then finalised through estimation of each dimension's internal consistency by reliability coefficients. Further validity evidence was also provided.

Results

Sample 1 - Pilot Study Procedures, Data Analyses, and Results

In an attempt to understand the structure of a group of measured variables and to find out latent constructs by removing unutilised measured variables and retaining utilised measured variables, exploratory factor analysis was conducted (Field, 2009). The FFPS was administered to 97 undergraduate students who have taken or are taking project-based courses. Before conducting exploratory factor analysis, the results of the parallel analysis, the Bartlett's test of sphericity and the Kaiser-Meyer-Olkin (KMO) measure of sample adequacy were examined in order to determine the appropriateness of factor analysis. Since the data were adequately normally distributed and variables were correlated, Maximum Likelihood as an extraction method (Fabrigar et al., 1999) and Direct oblimin as oblique rotation method (Costello & Osborne, 2005) were selected. For retention of the number of factors, instead of scree plot which can be problematic (O'connor, 2000) and may cause underestimating and/or overestimating (Zwick & Velicer, 1986), parallel analysis was used as recommended in the literature (Fabrigar et al., 1999; O'connor, 2000; Costello & Osborne, 2005).

Parallel analysis yielded three factors for FFPS. Accordingly, factors of each scale were extracted based on a fixed number of three factors. Bartlett's test was significant (BTS value= 1357.89, $p < 0.05$), indicating that for FFPS the identity matrix and correlation matrix were significantly different from each other. Moreover, the KMO measure of sampling adequacy of 0.90 revealed that it was appropriate to perform factor analysis (Tabachnick & Fidell, 2007). The three factors of FFPS were retained and interpreted as relating to development (DEV), understandability (UND),

and encouragement (ENC). The 57.47% of variance explained by the three factors were deemed conceptually sufficient (Scherer et al., 1988).

Items whose loading was above 0.30 were deemed satisfactory and therefore considered to load on a factor (Stevens, 2009; Tabachnick & Fidell, 2013). Item communalities were also inspected and items with communality below 0.40 were furthered scrutinised for inclusion or exclusion (Costello & Osborne, 2005). After the items were inspected based on the criteria mentioned, eight items were excluded from further analysis due to low correlation and seven items were removed because of cross-loading or no loading. As can be seen from Table 2, all the remaining items had pattern coefficients greater than 0.46. A few items with low communalities were observed, yet it was decided to retain them for further examination due to their significance to the three-factor model. Analysis of data from this pilot study culminated with FFPS consisting of 25 items and three dimensions. Table 2 illustrates factor loadings, communalities, internal consistency reliabilities, and factor correlation.

Three dimensions along with their definitions are as follows:

1) Developmental aspect (DA) (11 items): It is the degree to which formative feedback given to project work is perceived by team members (individuals and group) with characteristics such as giving direction during the process of project work revision, explaining how to revise in detail, and providing what needs to be done in order to improve weak performance etc.

2) Encouragement aspect (EA) (7 items): It is the degree to which formative feedback given to project work is perceived by team members (individuals and group) with characteristics such as motivation to revise, having a positive tone and manner, and showing the instructor cares about the work being performed etc.

3) Understandability aspect (UA) (7 items): It is the degree to which formative feedback given to project work is perceived by team members (individuals and group) with characteristics such as being easy to understand, consistent, easy to read, and has relevance to the topic/problem etc.

Table 2

Exploratory Factor Analysis with Oblique Rotation, Communalities and Reliability Analysis of all Items

	Factor loading					
Items	1	2	3	h^2	Cronbach's alpha if item deleted	Cronbach's alpha
Factor I: Development						0.92
Item 1	**0.92**	-0.09	-0.01	0.77	0.90	
Item 7	**0.86**	0.06	-0.08	0.72	0.90	
Item 3	**0.76**	-0.17	0.02	0.52	0.91	
Item 5	**0.69**	0.01	0.00	0.49	0.91	
Item 8	**0.69**	0.03	-0.03	0.47	0.91	
Item 6	**0.61**	0.08	0.14	0.54	0.91	
Item 9	**0.61**	0.22	0.00	0.53	0.91	
Item 11	**0.60**	0.15	0.09	0.52	0.91	
Item 10	**0.58**	0.16	0.04	0.46	0.91	
Item 2	**0.53**	0.24	0.03	0.47	0.91	
Item 4	**0.47**	0.13	0.19	0.43	0.92	
Factor II: Encouragement						0.88
Item 19	0.02	**0.84**	0.02	0.74	0.85	
Item 24	0.06	**0.82**	0.01	0.72	0.86	
Item 20	0.07	**0.64**	0.22	0.62	0.87	
Item 25	-0.03	**0.61**	0.13	0.43	0.87	
Item 22	0.17	**0.59**	-0.05	0.44	0.87	
Item 23	0.04	**0.59**	-0.03	0.35	0.88	
Item 21	0.19	**0.53**	0.14	0.50	0.87	
Factor III: Understandability						0.83
Item 12	-0.17	0.14	**0.84**	0.69	0.79	
Item 13	-0.06	0.08	**0.69**	0.48	0.81	
Item 15	-0.05	0.26	**0.56**	0.46	0.81	
Item 17	0.21	-0.27	**0.56**	0.38	0.82	
Item 18	0.22	-0.02	**0.48**	0.37	0.81	
Item 16	0.25	0.03	**0.47**	0.42	0.81	
Item 14	0.17	0.19	**0.47**	0.45	0.81	
Overall Cronbach's alpha						0.93
Factor correlations						
Factor 1	-					
Factor 2	0.41	-				
Factor 3	0.49	0.40	-			

Note: N=97. Boldface indicates highest factor loadings. Reliability represents composite Cronbach's alpha for items in a factor. Item reflects response choices of; 1=Never, 2=Rarely, 3=Occasionally, 4=Frequently, and 5=Very frequently. h^2=Communality after extraction.

Reliability analysis with Sample 1

In an attempt to measure internal consistency of the scale, Cronbach's alpha test was applied. While conducting the test, a separate Cronbach's alpha reliability coefficient was calculated for each of the scale's items with each dimension. As can be observed from Table 1, overall Cronbach's alpha reliability coefficient is 0.93,

indicating satisfactory and good internal consistency (DeVellis, 2003; Field, 2009). As for the internal consistency reliabilities of the items within each dimension, test results yielded 0.92 for DEV, 0.88 for ENC, and 0.83 for UND, which are regarded as showing good internal consistency as they are larger than 0.70, and are therefore accepted as a sign of acceptable reliability (Nunnally, 1978).

Sample 2 - Main Study Procedures, Data Analyses, and Results

The main study was conducted to confirm construct validity of the factor structure of the scores obtained from the 25-item FFPS, and to identify whether or not the number of factors fixed a priori. CFA was employed using the analysis of moment structures (AMOS version 20) statistical software package. The method of maximum likelihood estimations was employed and three aspects of FFPS were allowed to correlate to each other as seen in Figure 1, which illustrates the model specification and the parameter estimates.

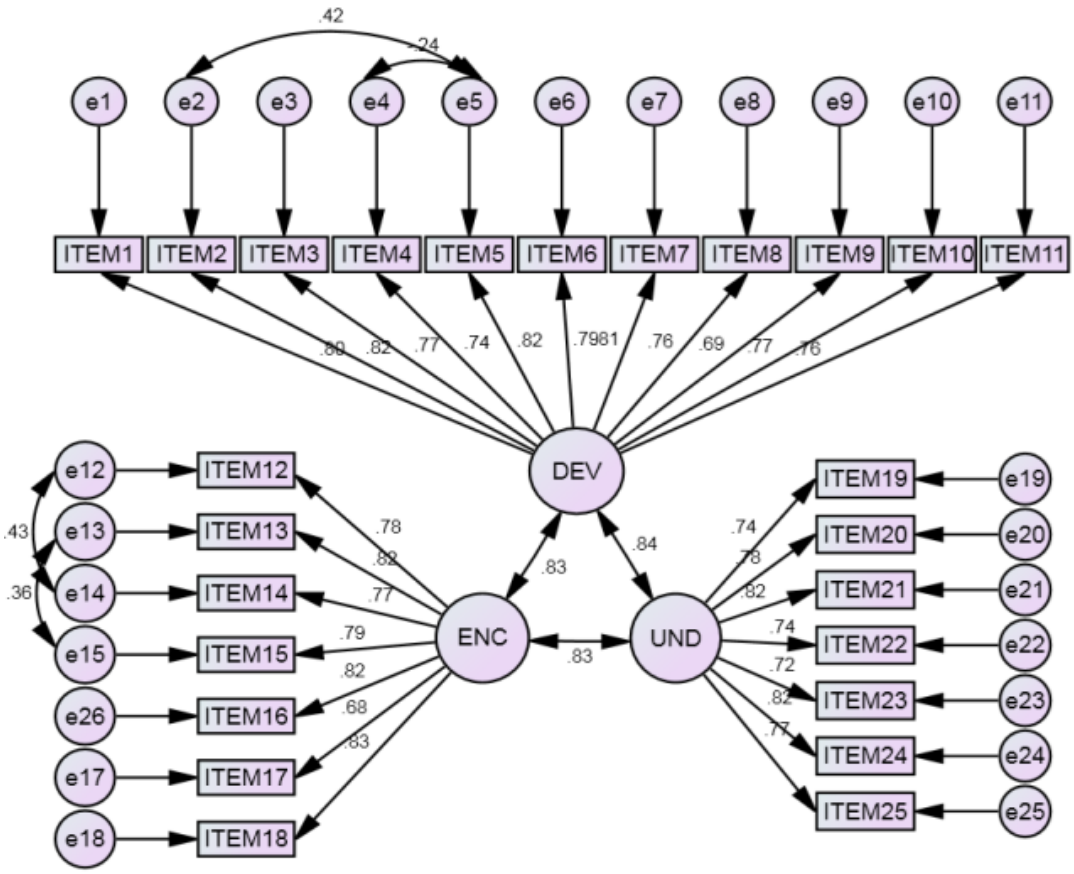

Figure 1

Standardized Coefficients for the Three-Factor Model of FFPS

Multiple goodness-of-fit indexes were used based on the suggestion of Brown (2006) in order to evaluate the overall model fit for FFPS. These are Chi-square ($\chi 2$) (Schermelleh-engel, Moosbrugger, & Müller, 2003), Comparative Fit Index (CFI;

Bentler, 1990), Tucker Lewis Index (TLI; Bentler, 1990), and Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993). Bentler (1990) reported that CFI and TLI in the range of 0.90-0.95 is indicative of an acceptable fit of the model. RMSEA equal to or smaller than 0.08 indicates acceptable fit (Browne & Cudeck, 1993; Schermelleh-engel, Moosbrugger, & Müller, 2003). Authors have criticised chi-square statistics due to being highly sensitive to sample size (Kline, 1998); therefore, it was elected not to be considered for this study.

CFA results suggested that the three-factor structure for FFPS has a reasonable fit to the data with all indices. Specification errors caused by shared variance among items were noticed, therefore, minor modifications to the model were deemed necessary in order to improve the fit. The modification indices between items guided to the covariate of e2-e5, e4-e5, e12-e14, and e13-e15. The final result suggested a CFI value of 0.92, a TLI value of 0.91, and RMSEA value of 0.078 which indicates an acceptable model fit. Moreover, factor loadings were found to be significant ($p < 0.001$). Figure 1 illustrates the model specification and the parameter estimates.

Convergent and discriminant validity with Sample 2

Convergent and discriminant validity are the tests used to control validity of the scale. Convergent validity refers to the degree to which each score of the item are correlated with the corresponding factor. According to Hair et al. (2010), three methods are indicators of convergent validity. The first and second indicators are that item loading under the corresponding factor and average variance extracted (AVE) for each factor should be greater than 0.50. As for last indicator, the construct reliability (CR) value should exceed 0.70. Results showed that item loadings changed between 0.68 and 0.83 and were significant in that it supports that all items are related to the corresponding factor. As shown in Table 3, AVE and CR value for each factor were larger than the suggested criterion (> 0.70). Together all the results supported the convergent validity of FFPS.

Table 3

Convergent Validity Results

	DEV	**ENC**	**UND**
AVE	0.603	0.621	0.595
CR	0.943	0.920	0.911

As for discriminant validity, it relates to the degree to which each factor measures a different dimension of the construct. The indicator, as suggested by Hair et al. (2010), is that Chi-square value for the default three-factor model and one-factor model should be significantly different. As indicated in Table 4, there is a significant difference between Chi-square values of the two models, suggesting that the three-factor model was better than the one-factor model.

Table 4

Model Comparison Results

Model 1 (three-factor)	**Model 2 (one-factor)**
χ^2 = 683.730	χ^2 = 1,016.747
sd = 268	sd = 271
p=.000	*p*=.000
$\chi^2_2 - \chi^2_1 = 333.017$	
$Sd_2 - sd_1 = 3$	

Discussion and Conclusion

The three-dimension Formative Feedback Perceptions Scale (FFPS) was developed following an in-depth literature review and semi structure interviews with participants. The current study was two-phase with the use of exploratory sequential mixed methods design. The first phase sought to explore the dimensions of formative feedback perceptions and then proceeded with a second phase in order to develop a questionnaire to measure formative feedback perceptions.

Findings from the qualitative phase of the study and review of the literature were used to develop the items and reveal the dimensions of FFPS. The interview transcripts were used to write the items and supported by statements that were used in relevant literature. Two independent sample groups of undergraduate students (Sample 1 and Sample 2) were used to validate FFPS through exploratory factor analysis and confirmatory factor analysis, which revealed a three-factor structure for the FFPS. Sample 1 consisted of 97 undergraduate students who participated in a pilot study in which exploratory factor analysis was used to test the factorial structure of the FFPS. In the main study, CFA was conducted to confirm the three-dimensional model of the FFPS on data gathered from Sample 2 (a further 250 undergraduate students). The 25-item FFPS was found to measure three dimensions of formative feedback perceptions named accordingly as follows with the corresponding items:

(1) Development Items: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
(2) Encouragement Items: 12, 13, 14, 15, 16, 17, 18
(3) Understandability Items: 19, 20, 21, 22, 23, 24, 25

The construct validity of the FFPS scores was ensured though CFA. Specifically, the results of the CFA indicated that the contribution of the items with high pattern coefficients to the factors under which they were located was significant. In addition, the study showed that the three-factor model was a reasonable fit to all indices (TLI=0.9; CFI=0.92; RMSEA=0.078). Further validation analysis tested for discriminant and convergent validity, and also showed that all the items converge on the corresponding dimension. Taken together, these results provide sufficient evidence that the construct validity of the FFPS scores were well established with the sample of undergraduate students. Examination of internal consistency estimates of the items by using Cronbach's alpha also demonstrated that all three dimensions (DEV=0.92; UND=0.83; ENC=0.88) and all the items (0.93) in the FFPS had high

internal consistency, indicating that FFPS generates reliable scores. Overall, it can be concluded that FFPS is a valid and reliable measure composed of three important dimensions essential to the investigation of formative feedback perceptions in project-related higher education courses.

The use of multiple approaches, specifically exploratory sequential design, can be considered a rigorous and systematic method because it enables the use of multiple forms of validity and reliability in order to qualitatively and quantitatively explore the multi-dimensions of formative feedback perceptions from the views of individuals in the population. It is not possible or sufficient just to present insights into formative feedback as a theoretical construct, by using either just qualitative or quantitative approaches. For that reason, the current study should prove particularly valuable as a means to reveal the dimensions of formative feedback perceptions, and to contribute a new, reliable, and validated research instrument to the collection of assessment and evaluation literature in higher education.

The growing body of research on the perceived value and effects of feedback in performance at the individual level has featured in the literature (Lizzio & Wilson, 2008; Poulos & Mahony, 2008; Price et al., 2010; Weaver, 2006). However, feedback given to individual members of the group or the group as a whole has not yet been extensively explored (Gabelica et al., 2012; London & Sessa, 2006), especially for project-based courses where educational products along with a project report and portfolio are produced by students as a team or group. Considering the operating of feedback in the complexity of group dynamics and the failing of conventional and standardised instruments to measure feedback on team-based projects from the students' point of view, this scale is a promising tool to be applied to explore and measure students' formative feedback perceptions/experiences on three dimensions, either to individual members or to the group as a whole.

It should be noted that the instrument proposed in the current study has a high applicability not only in science courses alone but also in engineering courses as well. Because, the way students are assessed during design and development of the project in undergraduate engineering courses is akin to the project in undergraduate science courses. Some exemplary engineering courses are documented in literature (Gibson, 2001; Razmov & Vlasseva, 2004) where the course is structured around the project accomplished by small group of students. As in this study, students in referred studies received feedback on their project and written reports in regular project meetings.

Future Research and Implication for Practice

The findings may have wider application to work-based and experiential learning which can also often involve group work. This would be an interesting avenue of further research. Moreover, given the difficulty of student-teacher interaction especially in large class settings, knowing student's perceptions on feedback might help tutors or instructors steer student learning in a way that

encourages deep approaches to learning. Previous studies highlighted the association of feedback preferences with surface and deep approaches to learning. Further studies are needed to understand how formative feedback perceptions inform or encourage these two learning approaches. In addition, future research is needed to validate and replicate a similar structure of the current scale with a more representative and larger sample from different disciplines, nations and cultures.

Limitations

Even though this study contributes to the existing knowledge on formative feedback by providing a new multidimensional instrument, there are also certain limitations that are worthy of mention. Firstly, selecting participants from the same department, albeit from four public universities, in order to reveal students' authentic experiences of formative feedback on project courses, may be a limitation to a certain extent in terms of ability to generalise the results of this study to other samples from other departments. In addition, self-reports by the participants during interviews and the completion of questionnaires may also be considered a limitation for the validity and reliability of the results of this study. Therefore, interpretation of the results of the current study should be considered based on these limitations.

Acknowledgments

This article is as a part of master's thesis supervised by Asst. Prof. Dr. Gülfidan Can in the Department of Computer Education and Instructional Technology at Middle East Technical University.

Appendix

Turkish and English Versions of the Formative Feedback Perception Scale (FFPS)

#	*Turkish*	*English*
	Genel olarak, projeme verilen yapılandırıcı dönütler....	**In general, formative feedback given to my project....**
DEV_1	… düzeltme sürecinde yol gösterir.	… gives direction during revision process.
DEV_2	… ayrıntılı olarak nasıl düzeltme yapacağımı açıklar.	… explain how to revise in detail.
DEV_3	… nasıl düzeltme yapacağım ile ilgili temel ipuçları içerir.	… includes basic tips about how to revise.
DEV_4	… düzeltme gereken yerleri açık şekilde gösterir.	… shows me clearly the place where revision is needed.
DEV_5	… performansın zayıf yönlerini geliştirmek için yapılması gerekenleri belirtir.	… provides what needs to be done to improve weak sides of performance.
DEV_6	… bakmam gereken yön hakkında beni yönlendirir.	… gives clues about which direction to look.
DEV_7	… etkilidir.	… is effective.
DEV_8	… yapıcıdır.	… is constructive.
DEV_9	… açıklayıcıdır.	… is well-explained.
DEV_10	… olumsuz noktaları sebepleriyle beraber verir.	… negative points are given with their justifications.
DEV_11	… gelecekteki projelerim için faydalıdır.	… helps me in future projects.
UND_1	… anlaması kolaydır / anlaşılırdır.	… is easy to understand.
UND_2	… okunması kolaydır (yazılı geri dönüt için).	… is easy to read (for written feedback).
UND_3	… düzeltmesi kolaydır / pratiktir.	… is easy to revise / practical.
UND_4	… tutarlıdır / çelişkili değildir.	… is consistent / not contradictory.
UND_5	… konuya ve soruna uygundur.	… is relevant to the topic and the problem.
UND_6	… performansın zayıf yönlerine dikkat çeker.	… draw attention to weak sides of performance.
UND_7	… önceden belirlenmiş değerlendirme ölçütlerini temel alır.	… is given based on the previously defined assessment criteria.
ENC_1	… yaptığım işe değer verildiğini hissettirir.	… shows that instructor cares about the work I have done.
ENC_2	… sarfettiğim emeği dikkate alır.	… recognizes the effort I have made.
ENC_3	… düzeltmeler için beni teşvik eder.	… motivates me to revise.
ENC_4	… çoğunlukla olumludur.	… is mostly positive.
ENC_5	… olumsuz şeyleri olumlu şekilde sunar.	… presents negative things in a positive way.
ENC_6	… tonu ve yaklaşımı olumludur.	… has positive tone and manner.
ENC_7	… olumlu ile eleştirel arası dengelidir.	… has a balance between critical and positive.

* **DEV:** Development; **UND:** Understandability; **UNC:** Encouragement.

References

Bandura, A. (1991). Social Cognitive Theory of Self-Regulation. *Organizational Behavior and Human Decision Processes*, *50*, 248–287.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74. https://doi.org/10.1080/0969595980050102

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education*, *38*(6), 698–712. https://doi.org/10.1080/02602938.2012.691462

Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research* (1st ed.). New York, NY: Guilford Publications.

Browne, M. W., & Cudeck, R. (1993). Alternative Ways of Assessing Model Fit. In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models* (pp. 136–162). Beverly Hills, CA: Sage.

Clark, I. (2012). Formative Assessment: Assessment Is for Self-regulated Learning. *Educational Psychology Review*, *24*(2), 205–249. https://doi.org/10.1007/s10648-011-9191-6

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 & Evaluation*, *10*(7), 1–9.

DeVellis, R. F. (2012). *Scale development: Theory and applications.* (2nd ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Dillman, D. A. (2007). *Mail and Internet surveys: The tailored design method* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Draper, S. W. (2009). What are learners actually regulating when given feedback? *British Journal of Educational Technology*, *40*(2), 306–315. https://doi.org/10.1111/j.1467-8535.2008.00930.x

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

Fernandes, S., Flores, M. A., & Lima, R. M. (2012). Students' views of assessment in project-led engineering education: findings from a case study in Portugal. *Assessment & Evaluation in Higher Education*, *37*(2), 163–178. https://doi.org/10.1080/02602938.2010.515015

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London, UK: SAGE

Publications.

Gabelica, C., Bossche, P. Van Den, Segers, M., & Gijselaers, W. (2012). Feedback, a powerful lever in teams: A review. *Educational Research Review*, *7*(2), 123–144. https://doi.org/10.1016/j.edurev.2011.11.003

Gibson, I. S. (2001). Group Project Work in Engineering DesignÐLearning Goals and their Assessment *. *International Journal of Engineering Education*, *17*(3), 261–266.

Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). John Wiley & Sons.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis* (7th ed.). NJ: Prentice-Hall.

Hattie, J. (2008). *Visible learning: A synthesis of over 800 Meta-analyses relating to achievement*. London, UK: Routledge.

Helle, L., Tynjälä, P., & Olkinuora, E. (2006). Project-Based Learning in Post-Secondary Education – Theory, Practice and Rubber Sling Shots. *Higher Education*, *51*(2), 287–314. https://doi.org/10.1007/s10734-004-6386-5

Higgins, R., Hartley, P., & Skelton, A. (2001). Getting the Message Across: The problem of communicating assessment feedback. *Teaching in Higher Education*, *6*(2), 269–274. https://doi.org/10.1080/13562510120045230

Hounsell, D., McCune, V., Hounsell, J., & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development*, *27*(1), 55–67. https://doi.org/10.1080/07294360701658765

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: Guildford.

Knight, P., & Yorke, M. (2003). *Assessment, Learning and Employability*. Maidenhead, UK: SRHE/Open University Press.

Lizzio, A., & Wilson, K. (2008). Feedback on assessment: students' perceptions of quality and effectiveness. *Assessment & Evaluation in Higher Education*, *33*(3), 263–275. https://doi.org/10.1080/02602930701292548

London, M., & Sessa, V. I. (2006). Group Feedback for Continuous Learning. *Human Resource Development Review*, *5*(3), 303–329. https://doi.org/10.1177/1534484306290226

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. https://doi.org/10.1080/03075070600572090

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

O'connor, B. P. (2000). SPSS and SAS programs for determining the number of components

using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396–402. https://doi.org/10.3758/BF03200807

Oppenheim, A. N. (1992). *Questionnaire Design, Interviewing and Attitude Measurement*. New York: Bloomsbury Publishing.

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

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback : all that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, *35*(3), 277–289. https://doi.org/10.1080/02602930903541007

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, *28*(1), 4–13. https://doi.org/10.1002/bs.3830280103

Razmov, V., & Vlasseva, S. (2004). Feedback Techniques for Project-based Courses.

Rowe, A. D., & Wood, L. N. (2008). Student Perceptions and Preferences for Feedback. *Asian Social Science*, *4*(3), 78–88. https://doi.org/10.5539/ass.v4n3p78

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144. https://doi.org/10.1007/BF00117714

Sadler, D. R. (1998). Formative Assessment : revisiting the territory Formative Assessment : revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 77–84.

Scherer, R. F., Wiebe, F. A., Luther, D. C., & Adams, J. S. (1988). Dimensionality of coping: factor stability using the Ways of Coping Questionnaire. *Psychological Reports*, *62*(3), 763–70. https://doi.org/10.2466/pr0.1988.62.3.763

Schermelleh-engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models : Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online*, *8*(2), 23–74.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting Self-Regulation in Science Education: Metacognition as Part of a Broader Perspective on Learning. *Research in Science Education*, *36*(1–2), 111–139. https://doi.org/10.1007/s11165-005-3917-8

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

Stevens, J. P. (2009). *Applied Multivariate Statistics for the Social Sciences* (5th ed.). New York: Routledge/Taylor & Francis.

Tabachnick, B., & Fidell, L. (2007). *Using multivariate statistics*. Boston: Allyn & Bacon.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (3rd ed.). New York, NY: Pearson.

Taras, M. (2002). Using Assessment for Learning and Learning from Assessment. *Assessment & Evaluation in Higher Education*, *27*(6), 501–510. https://doi.org/10.1080/0260293022000020273

Taras, M. (2005). Assessment – Summative and Formative – Some Theoretical Reflections. *British Journal of Educational Studies*, *53*(4), 466–478. https://doi.org/10.1111/j.1467-8527.2005.00307.x

Thomas, J. W. (2000). *A Review of Research on Project-Based Learning*. San Rafael, CA: Autodesk Foundation.

Weaver, M. R. (2006). Do students value feedback? Student perceptions of tutors' written responses. *Assessment & Evaluation in Higher Education*, *31*(3), 379–394. https://doi.org/10.1080/02602930500353061

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3–14. https://doi.org/10.1016/j.stueduc.2011.03.001

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