Developing Questionnaires as a Tool of Collecting Information for Program Evaluation in Medical Education

Kemal Turhan[a,1](kturhan.tr@gmail.com)
Burçin Kurt[1](burcinnkurt@gmail.com)
Aslı Yazağan[b,2](ayazagan@gmail.com)
Uğur Toprak[1] (toprakugur@ktu.edu.tr)
Serbülent Ünsal[1] (serbulentu@gmail.com)

[a]*Karadeniz Technical University, School of Medicine, Trabzon, Turkey*
[b]*Recep Tayyip Erdoğan University, Ardeşen Vocational School, Rize, Turkey*

Abstract– This study presents a student questionnaire development process, experiences gained, and results. A six-step questionnaire development process was followed: literature review, regular stakeholder meetings, initial item generation, preliminary survey, item reliability analysis, and factor analysis.

All developed questionnaires were qualified using statistical methods. The number of items in the questionnaires was reduced using factor analysis. After the proposal has been transformed into a software project, we have been completed the medical student feedback system(MSFS) that is very easy to use, flexible, compatible with mobile devices and conforms to scientific requirements for program evaluation.

A framework including student and faculty interactivity; non-time-consuming, efficient feedback collection; and scalable, satisfactory data for program evaluation was established. MSFS has also been integrated with social media applications to make real-time interaction with students. The framework presented in this study can be used when it is needed to develop effective student feedback system.

Keywords– Medical Education, Program Evaluation, Data Reduction, Student Feedback, Questionnaire development

---

[1]*Corresponding Author*

1. Introduction

Due to the rapidly changing nature of medical knowledge, medical education program development and management must continuously evolve and program effectiveness must be continually assessed. In medical education research, program evaluation has become important. Evaluation of education programs is required by medical education accreditation organizations (LCME 2015). Program evaluation is defined as (CDC 2015) the systematic collection and analysis of information related to the design, implementation, and outcomes of a program, for the purpose of monitoring and improving the quality and effectiveness of the program. To address the question of how a program works, not only curricular content and program outcomes, but also program processes must be considered (Haji, Morin, and Parker. 2013).

Student feedback is one of the important input data sources for all program evaluation models, and it has vital role in faculty-student interaction. Faculty-student interaction refers to the degree of responsiveness, helpfulness, concern, approachability, compassion, and understanding that instructors display towards their students (Paolini 2015). Faculty-student interaction outside of the classroom is an important influence on student success and retention. Students' feedback plays a primary role as a communication tool regarding the program (Poulosa, and Mahon 2008). Researchers suggest implementing an effective and operative feedback system to provide continuous communication between students and faculty (Garner, Gusberg, and Kim 2014).

Feedback can be provided through questionnaire surveys (Chan, Konrad, Gonzalez, Peters, and Ressa. 2014). Questionnaires can be useful for collecting information about behaviours that are not directly observable (Artino, La Rochelle, Dezee, and Gehlbach 2014). It is assumed that students are willing to complete questionnaires. However, this assumption is not usually correct. Filling questionnaires is considered time-consuming, and students believe that their evaluations will not affect the education program. While the effectiveness of feedback is key for accurate program evaluation, little research has focused on students' perceptions of feedback and the process of contributing feedback. Elucidate feedback experiences with a narrative interviewing approach and report the negative experiences of students during the feedback collection process (Poulosa, and Mahon 2008; Urquhart, Rees, and Ker 2014). It was also reported that students admire feedback when it is clear and instructive (Small, and Attree 2016).

In this study, we conducted meetings with students, student affairs staff, and teachers to develop a preliminary questionnaire. We tried to eliminate the negative beliefs of students by including them in the questionnaire development process. Because, students and teachers believe in usefulness of questionnaires, but they do not get serious enough for some reasons (Richardson, 2005). Moreover, the number of items in the questionnaire was reduced to provide a satisfactory experience for students.

2. Material and methods

The integrated medical education system was adopted six years ago by the Medical Faculty of Karadeniz Technical University. In this system, the curriculum is designed considering human body systems. In the first three years, there are core lessons designed for minimal

clinical exposure. The fourth and fifth years comprise a mixture of clinical clerkship and medical courses. The last year is the internship period in the hospital.

To pass each year, students need to achieve a certain score determined by the educational committee assessment. Every year has two semesters, and there are six, seven, and nine committee assessment in the first, second, and third years, respectively. At the end of each committee assessment, students must take examinations and, in each semester, they must take a final examination. In the fourth and fifth years, an oral examination and a written examination are conducted at the end of clinical clerkships. Additionally, various formative examinations that affect the main mandatory examination may be conducted in different committee assessments and clinical clerkship periods in every year.

The National Medical Education Accreditation Institute (UTEAK) (Frye, and Hemmer 2012) establishes the National Standards for Undergraduate Medical Education. Medical schools wishing to become accredited in Turkey apply to UTEAK. According to UTEAK standards, faculties of medicine must establish a program evaluation system that monitors the curriculum achievements and student progress by gathering the opinions and proposals of stakeholders about the ongoing education program. In the process of accreditation in the Medical Faculty of Karadeniz Technical University, various working groups were established, such as a curriculum development group, a measurement and assessment group, and a program evaluation group.

Program evaluation should be designed to detect changes that are fundamental for education programs. Change can occur at any stage and can be intended or unintended, and directly observed or non-observed. The education program evaluation should consider all these issues. Small changes in program elements may lead to large changes in outcomes (LCME 2015). It is important to address not only whether the program worked, but also how change has occurred (Haji, Morin, and Parker 2013).

There are commonly used program evaluation models such as the Kirkpatrick four-level model, Logic Model and CIPP Model. Any of these models could address all program evaluation issues. Each has its own advantages. The CIPP model was mostly utilized in the program evaluation studies during the accreditation process of the Medical Faculty of Karadeniz Technical University (UTEAK 2015).

UTEAK program evaluation standards studies were undertaken by the Program Evaluation Working Group (PEWG). PEWG was composed of instructors from various specialties, students in different school years, and statisticians. As part of the program evaluation effort, questionnaire development (QD) studies were conducted to retrieve effective and satisfactory student feedback. PEWG decided to use surveys as one of the important data sources mentioned in the CIPP Model.

Artino et al. present a guide on developing questionnaires for education research (Artino, La Rochelle, Dezee, and Gehlbach. 2014), and PEWG implemented similar steps. The scheme of the proposed six-step questionnaire development process is given in Graph 1.
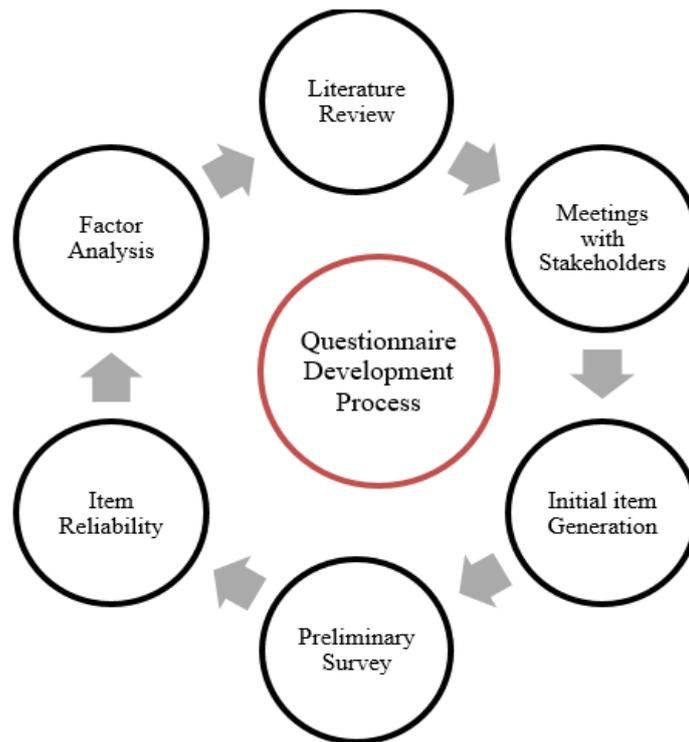
Graph-1: *Six-step questionnaire development process*

2.1. Literature review

UTEAK standards and literature suggest that student attitudes about the feedback collection process are affected by instructors' personalities, especially leadership skills, sociability, objectiveness, and supportiveness (Paolini, 2015; Urquhart, Rees, and Ker 2014). Instructors' personalities and communication with students were considered while selecting PEWG members. Moreover, we decided to work with as many students as possible in the QD Process. Each year, we had five student members in PEWG meetings. We invited more students in addition to PEWG members according to the meeting agenda.

2.2. PEWG Meetings with Stakeholders

PEWG arranged 50 meetings during 2012-2015. In the first meetings, a literature review on program evaluation and survey tools was conducted. Questionnaire items were discussed by group members and faculty for nine months. Questionnaires from other medical faculty members were considered, and items were reviewed. After deciding questionnaire items, the implementation method of the questionnaire was discussed with stakeholders such as semester coordinators, student affairs, PEWG members, and invited students. Teachers who visited other universities shared their observations about education program evaluation studies there. Evaluation reports were prepared and presented to the faculty administrators and other related working groups. To raise awareness among students about the importance of program evaluation, informative emails were sent, brochures were displayed on faculty boards, and feedback boxes for quick feedback were placed in visible locations near teaching classes.

2.3. Initial item generation

Questionnaire items were developed in interactive meetings conducted by PEWG. Comparison of numbers of questionnaire items is given in Table 1.

Table-1: *Comparison results of number of questionnaire items and reduced number of questionnaire items*

Questionnaire	Number of Items	Reduced Number of Items	Number of Students
Evaluation questionnaire for first, second and third year teaching committees (QC)	14	7	1013
Evaluation questionnaire for committee instructors (QCI)	11	4	2357
Evaluation questionnaire for committee exams (QCE)	8	8	1013
Evaluation questionnaire for clinical clerkship (QCC)	10	6	1650
Evaluation questionnaire for clinical clerkship instructor (QCCI)	17	6	7192
Evaluation questionnaire for clinical clerkship exams (QCCE)	9	9	1650

2.4. Preliminary Survey

After comprehensive research over several months, a preliminary survey was conducted on students under the supervision of instructors. A feedback report was created based on survey data and presented to the faculty administrator.

2.5. Item Reliability Analyses

After the preliminary questionnaire, item reliability analyses were performed using statistical techniques Cronbach's Alpha. Since Cronbach's Alpha values did not change positively and significantly, PEWG did not reduce the number of items in the questionnaire after item reliability analysis. Cronbach's Alpha values can be seen in Table 1. If we had looked at the Cronbach's Alpha values and reduced the number of items, it could have been caused a loss of information about validity of questionnaire. Because Cronbach's Alpha mostly related about reliability. If the Cronbach's Alpha value was too low, we could have made changes to the questions.

2.6. Factor Analysis for Item Reduction

After the preliminary survey, we inferred that students had more negative than positive experiences in the feedback collection process. Students and instructors reported complaints about the large number of items and similar repetitive questions. To improve students' perceptions of the process, PEWG decided to reduce the number of items.

Factor analysis was used as the data reduction method. A questionnaire needs to be validated to verify that it measures what it is intended to measure (Field 2005). Factor analysis attempts to identify the underlying variables or factors that explain the pattern of

correlations within a set of observed variables. It is often used to identify a small number of factors that explain most of the variance observed in a much larger number of variables (UCLA, 2015). Table 2 shows the factor analysis results. Items were clustered into five subgroups using correlation between components and items. Items in each subgroup explain a component of the questionnaire. Items in a group have a very similar meaning. Interestingly, we did not notice this semantic closeness before the factor analysis (Table 2).

Table-2: *Factor analysis results for qc: probability relation between items and rotated components* [2]

	1	2	3	4	5
Component Number	1	2	3	4	5
Cumulative % variance explained	14,1	27,1	40,1	50,5	59,0
% of Variance explained	14,1	13,0	13,0	10,4	8,5
The practical study opportunity was sufficient.	,752				
Practical study hours were sufficient.	,744				
Theoretical classes were related to practical studies.	,671				
The committee period was sufficient for education.		,735			
I did not have difficulties integrating what I learned in the committee with what I had learned earlier.		,595			
Learning objectives set at the beginning of the committee had been achieved.		,571			
Subjects had been sequenced to ensure the integrity of knowledge and to facilitate learning.		,542			
I attended the classes.			,765		
The committee encouraged the attendance of classes.			,697		
I had the opportunity to actively participate in the classes.			,623		
Committee goals and learning objectives have been announced.				,797	
Goals and learning objectives have been announced in practical studies.				,599	
The course schedule was followed.					,841
I had to be prepared for the classes.					,750

---

[2] ExtractionMethod: Principal Component Analysis. RotationMethod: VarimaxwithKaiserNormalization

Considering these results, some questions were selected directly, some were combined into one question, and the number of items in the questionnaire was reduced to seven questions. The items are:

1.   The subjects were sequenced to ensure the integrity of knowledge and to facilitate learning.
2.   The course schedule was followed.
3.   The committee period was sufficient for education.
4.   The practical studies were sufficient.
5.   I attended the classes.
6.   I had to be prepared for the classes.
7.   I had the opportunity to actively participate in the classes.

3. Results

Surveys are important data sources for program evaluation, and the Medical Faculty of Karadeniz Technical University decided to design student feedback questionnaires for the accreditation process. Feedback questionnaires could be developed for the overall education program or a specific education module. Alotaibi et al. developed a questionnaire solely for a respiratory care education program(Alotaibi, and Youssef 2013). In our study, we developed questionnaires for overall education program evaluation. As noted in other studies (UTEAK, 2015; Alotaibi, and Youssef 2013), problems arise because of the frequency of conducting surveys and the number of items in the surveys. To help resolve these problems, a six-step questionnaire development process was implemented in this study: literature review, meetings with stakeholders, initial item generation, preliminary survey, item reliability analysis, and factor analysis (Figure 1). We developed six different questionnaire systems for four different educational contexts (Table 3): General Teaching Quality, General Clinical Clerkship Quality, Instructional Quality in Teaching and Clinical Clerkship, and Exam Quality in Teaching and Clinical Clerkship.

Table-3: Questionnaire contexts

Purpose	Developed Questionnaire
General Teaching Quality	Evaluation questionnaire for 1,2,3 year teaching committees (QC)
General Clinical Clerkship Quality	Evaluation questionnaire for clinical clerkship (QCC)
Exam Quality	Evaluation questionnaire for clinical clerkship exams (QCCE)
	Evaluation questionnaire for committee exams (QCE)
Instructional Quality	Evaluation questionnaire for clinical clerkship instructor (QCCI)
	Evaluation questionnaire for committee instructors (QCI)

QC was developed for General Teaching Quality. QCC was developed for General Clinical Clerkship Quality. QCI and QCCI were developed for Instructional Quality. QCE and QCCE were developed for Exam quality.

At the end of the questionnaire development process, students took a paper version of the preliminary questionnaire under the supervision of instructors. The paper-based survey provided more interaction between faculty and students, and enabled easier management of the process. At the end of the QD process, we obtained a stable, scalable, validated, and verified questionnaire system. Thus, the requirement analysis for the web-based survey system was completed before implementing it.

4. Discussion

Some studies emphasized that student feedback is vital, especially for educational quality management efforts. Nevertheless, the fact that this processes are not being conducted well (UCLA 2015). They suggested that it is required an effective and efficient student feedback system (UCLA 2015). It was emphasized that students should be in the mood of to fill a questionnaire for effectivity and efficiency of feedback in the study (Blair 2014).

This study presents our extensive QD process for program evaluation experience that can be shared. In our study, the broad participation of students and other stakeholders was important. Students are at the centre of medical education programs, and the main outcomes of education programs are behavioural, attitudinal, and informational changes in the students. Other studies (Donner-Banzhoff, Merle, Baum, and Basler 2003; Sostok, Coberly, and Rouan 2002) conducted the questionnaire development process including only certain research groups with limited stakeholder participation. In some studies, (UTEAK 2015; Alotaibi, and Youssef 2013), data sources were limited to the educational environment but in our study, data from different educational contexts, such as experiential information from other medical schools, was included as a part of the development process. We gained valuable experience in retrieving feedback as well as eliciting the broad participation of students and stakeholders.

Questionnaire is not the only way to collect information for program evaluation. There are many methods such as qualitative interviews, focus groups interviews, document review that has some advantages and disadvantages (NCS 2017).
Although they are valuable information collecting methods, they have some disadvantages:

- Trained and skilled staff requirements,
- They can be time consuming,
- It is usually difficult and complex to evaluate collected information (Mennin 2010).

This of course does not mean that other methods should not be used. The data collection methods should also be used in the program evaluation efforts for detecting all changes in medical education progress.

5. Conclusion

Following student feedback about the preliminary questionnaire, many questionnaire items were deleted. All developed questionnaires were qualified using statistical methods. The number of items in the questionnaires was determined using the factor analysis technique. Factor analysis enabled us to obtain statistically maximum variance with minimum number of items in the questionnaires. Five-to-seven point odd-number scales are usually used in surveys, but the four-point Likert scale (strongly disagree, disagree, agree, and strongly agree) was used in our QD process to facilitate the impartial selection of students.

We had planned to migrate the framework established during QD to the web-based system. Thus, the requirement analysis for the web-based survey system was completed during the QD process.

The envisioned online questionnaire has contain a page informing students about their feedback process. This system will help to eliminate students' negative belief that feedback will not be considered. Other advantages of the web-based system are easy analyses of questionnaire results and updates to the feedback system.
At the end of the QD development effort, we accomplished a framework that enabled student-faculty interactivity, efficient feedback collection, and a scalable system for program evaluation.

Feedback systems should be 1) easily applicable; 2) improved and verified with statistical techniques; 3) scalable with optimal application frequency and an optimal number of items, and adaptable to quick changes; and 4) responsive in order to enable traceability by stakeholders.

Conforming with these four the requirements, MSFS has also been integrated with social media applications such as Instagram and Facebook to make notifications for surveys and to give information about what actions are taken for student feedbacks.
As a result, feedback systems and QD efforts should be planned as long term projects utilizing technology and including broad stakeholders' participation.

In this study, we tried to develop and implement a method for collecting systematic and continuous information for program evaluation by minimizing the disadvantages and extending advantages of questionaries' as a method of collecting data.

References

.

Alotaibi G, Youssef A. Development of an assessment tool to measure students' perceptions of respiratory care education programs: Item generation, item reduction, and preliminary validation. J Fam Comm Med. 2013; 20(2): 116-122.

Artino AR, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research. Med Teach. 2014; 36: 463-474.

Chan PE, Konrad M, Gonzalez V, Peters MT and Ressa VA. The Critical Role of Feedback in Formative Instructional Practices. IntervSch Clinic. 2014; 50(2): 96-104.

Donner-Banzhoff N, Merle H, Baum E, Basler HD. Feedback for general practice trainers: developing and testing a standardised instrument using the importance quality score method. Med Educ. 2003; 37: 772-777.

Factor   Analysis.   UCLA: Statistical   Consulting   Group.   Available   at: http://www.ats.ucla.edu/stat/spss/output/factor1.htm. Accessed September 2, 2015.

Field, A. P. Discovering statistics using SPSS. 2nd edition. London: Sage, 2005.

Frye AW, Hemmer PA. Program evaluation models and related theories: AMEE Guide No. 67. Med Teach. 2012; 34: 288-299.

Function and Structure of a Medical School 2015. Available at: http://lcme.org/publications/. Accessed September 6, 2015.

Garner MS, Gusberg RJ, Kim AW. The Positive Effect of Immediate Feedback on Medical Student Education During the Surgical Internship. J Surg. 2014; 71: 391-397.

Haji F, M. Morin M, Parker K. Rethinking program evaluation in health professions education: beyond "did it work?". Med Educ. 2013; 47:342-351

Introduction to program evaluation for public health programs: A self-study guide. Available at: http://www.cdc.gov/eval/guide/. Accessed September 3, 2015.

Mennin S. Complexity and health professions education: a basic glossary. J EvalClinPract. 2010; 16: 835-837.

National   Standards   for   Undergraduate   Medical   Education   2011.   Available   at: http://www.uteak.org.tr/condensed_english_version/14. Accessed September 5, 2015.

Paolini A. Enhancing Teaching Effectiveness and Student Learning Outcomes. J Eff Teach. 2015; 15(1): 20-33

Poulosa A, Mahon MJ. Effectiveness of feedback: the students' perspective. Assess Eval Higher Educ. 2008; 33(2): 143-154.

Sostok MA, Coberly L, Rouan G. Feedback Process between Faculty and Students. Acad Med. 2002; 77(3): 267.

Urquhart LM, Rees CE, Ker JS. Making sense of feedback experiences: a multi-school study of medical students' narratives. Med Educ. 2014; 48(2): 198-203.