Use of Reaction Evaluation to Improve E-learning Programs: Case of K University

SUK HWANG

Human Resources Development Institute, (330-240) Korea University Technology and Education,
Budae-Dong 415-1, Cheonan Chungnam, Korea

shwang@kut.ac.kr

Abstract

Many e-learning related educational institutions implement reaction evaluation at the end of the e-learning programs. As with off-line programs the reaction evaluation of the e-learning programs measure the perception of students and trainees about the various aspects of the programs. Although the reaction evaluation is used widely, the effectiveness of it is in question. To be an effective reaction evaluation the purpose of the evaluation, design of the questionnaire, and the use of the findings should be adjusted consistently.

K university has used a few hundred e-learning programs over the last 6 years and utilized the reaction evaluation as each program was completed. This study examined how categories and the number of questions for each category has been modified over the years. After examining the major changes of the reaction evaluation this study suggested ways to improve the reaction evaluation of e-learning programs.

Introduction

According to Kirkpatrick, evaluation of training programs has four levels. Among the four levels of evaluation, the first level evaluation which is called reaction evaluation is the most widely used level due to its convenience of collecting information (Sugrue & Kim, 2004). It measures trainees' perception on the various aspects and important issues of the training program. Main purpose of the reaction evaluation is to improve the programs by applying the findings of the reaction evaluation. Reaction evaluation should meet the needs of the program and the organization to which the program is belong (Dixon, 1996). However, the effectiveness of the reaction evaluation is turned out to be very low, working as a rite of passage of the training programs (Tannenbaum & Yukl, 1992).

For its convenience almost all types of training programs use reaction evaluation. As with the other training programs most of the e-learning programs use the reaction evaluation because of the similar aspects of the training and common evaluation items. Thus the procedures and standards for reaction evaluation of e-learning programs are likely to be very similar with that of offline training programs. In the same context the problems with effectiveness of the reaction evaluation of the e-learning programs could turn out to be similar with that of the off-line training programs. Among the many ways to measure the effectiveness of the reaction evaluation one of major factor for ensuring its effectiveness is alignment of purpose of the reaction evaluation with design of questionnaire and with the use of the evaluation findings.

K university has developed and utilized a few hundred e-learning programs for its students and other trainees over the last 6 years. Main delivery method of the e-learning program is the blend of on-and-off-line sessions, which uses e-learning contents for some period of time followed by off-line practices and experiment sessions. As the students and trainees completed the e-learning programs they are required to fill out the reaction evaluation questionnaire. It evaluates various aspects of the e-learning program including overall satisfaction, participation, quality of the contents, simulator, synchronous discussion, functionality, operation personnel, lecturer and assistants, and additional functions. The university has reported the findings annually and accumulated the results and findings of the reaction evaluation questionnaires.

The purpose of this study is to find the ways to improve the effectiveness of the reaction evaluation of e-learning programs, using a case of K university. It assumes that the effective reaction evaluation contributes to the improvement of e-learning programs. This study looks at the major aspects of the reaction evaluation of the case and examines us-

age of the findings of the evaluation. After discussing the problems with the usage it suggests ways to be an effective reaction evaluation which contribute to the improvement of the e-learning programs.

Reaction evaluation of e-learning programs of K University

K university relies on the use of multimedia contents for most of e-learning programs. For the university's e-learning programs the instructional strategy of contents and delivery methods were regarded as major factors affecting educational quality and achievements of the students. The multi-media contents are made of 16 elements, including course introduction, reminder of prerequisites, scenario, objective, check list for prerequisite, introduction of lesson, use of graph and table, further study cue, quiz, dictionary, simulation, summary, evaluation, information on related studies, and unit satisfaction question. Although the instructors are allowed to modify the portion of learning activity, once the multimedia contents are introduced in the program they are regarded as a predetermined material which are to be mastered in advance before starting hands-on experiment and evaluation. The instructors let the student use the contents for certain period of time, then continues to do laboratory or practice session and other interactive activities. All learning activities are managed by the instructors with the help from the technical support personnel. As interactive activities synchronous discussions and quizzes were included in the e-learning programs. The grade of students and trainees is based on the level of participation in the discussion, quality of homework, and progress of the study using contents. When the students and trainees completed the on-line portion of blended programs, they were required to fill out the reaction evaluation questionnaire.

Off-line reaction evaluation questionnaire looks into program purpose and contents, material, instructional strategy, media, instructor, instructional activity, length of the program, facility, administrative support, overall evaluation, and suggestion which measures effectiveness of almost all aspects of training (Lee & Pershing, 2002). Compared to the off-line training programs major aspects of e-learning programs are different. Items for off-line reaction evaluation need to be modified to fit to the major aspects of the e-learning programs when we plan e-learning reaction evaluation. K university developed categories for reaction evaluation questionnaire of e-learning programs after considering major components and procedural stages of its e-learning programs which included level of participation, contents, simulator, synchronous discussion, functionality, technical support personnel, lecturer and assistants, and other supporting functions as shown in Table 1.

Table 1. Categories and Description of the Reaction Evaluation

Components	**Description**	**Number of Questions**
Level of participation	Self-rated level of student participation in learning activities, follow the notices	2
Contents	Objectives, lesson contents, supporting material,	6
Simulator	Reading manual, ease of use, effectiveness of the online practices	5
Synchronous discussion	Facilitation of discussion, facilitating role of the instructor	3
Functionality	Downloading time, manifestation of functionality	5
Technical support personnel	Effectiveness of the help from the technical operating personnel	3
Lecturer and assistants	Responsiveness, discussion, assistance, feedback,	8
Other supporting function	Use of e-library, community, help function	4

Method

This study reviews the results of the reaction evaluation of e-learning programs over the last 6 years. After analyzing the data frequencies and mean scores were tabulated to determine the level of reaction for each category. Due to the changes of some of the categories since 2006 some questions of the new categories have been re-classified to fit to the previous categories.

Results

This study calculated the average scores for each category of the reaction evaluation questionnaire over the past 6 years as shown in Table 2. Integrated mean scores for categories are over 4.0 except 'other support' category.

Table 2. Average Scores over the 6 Years

Category	*Contents*	*Practice*	*Discussion*	*System*	*Operating support*	*Instructor*	*Other support*
Mean	4.08	4.09	4.03	4.04	4.09	4.05	3.17

To look into the detailed information of each category the average scores of the categories in each year since 2003 were tabulated as shown in Table 3.

Table 3. Scores for Each Semester

Semester/ category	*Contents*	*Practice*	*Discussion*	*Systemic maintenance*	*technical support*	*instructor*	*Other support*
2nd 2003	3.71	3.70	3.57	3.87	3.46	3.76	2.93
1 st 2004	4.19	3.86	3.47	3.97	3.97	3.68	3.08
2nd 2004	4.17	4.03	3.62	3.77	4.02	3.60	3.03
1 st 2005	4.12	3.97	4.16	4.10	3.91	4.12	3.61
2nd 2005	4.13	3.97	4.15	4.18	3.97	4.14	3.68
1 st 2006	4.28	4.34	N/A	N/A	4.07	N/A	3.09
2nd 2006	4.03	N/A	4.08	N/A	4.18	4.29	2.79
1 st 2007	4.12	4.3	4.3	4.1	4.23	4.32	N/A
2nd 2007	4.12	4.3	4.4	4.2	4.43	4.30	N/A
1 st 2008	4.02	4.2	4.3	4.1	4.33	4.17	N/A
2nd 2008	4.00	4.2	4.3	4.1	4.43	4.17	N/A

Noticeable change in the mean scores is shown in 2006. In spring of 2006, the scores for three categories were not shown, followed by two missing categories in fall of the same year. 'Other support' which had been rated as the lowest among the categories disappeared from the year of 2007.

Examination of the questionnaire over the years revealed that there were some modifications of questionnaires after the 3rd year of reaction evaluation. It was also found that, from the year of 2006, the number of questions was decreased for the purpose of expediting the evaluation procedure. After reviewing all new and modified questions of the questionnaire it was possible to allocate the questions to the already established categories. This study then looked at how the questions were modified in each category. As this study looks at the dropped questions it showed that the questions with low scores were dropped as shown in Table 4.

Table 4. Change of the Questions

Category	*Change of questions*	*Number of questions remained*
Participation	Category deleted	0
Contents	Remained	5
Simulator	Category deleted	0
Synchronous discussion	Replaced with a question of 'role of the instructor in the discussion'	1
Functionality	Changed to infrastructure maintenance	1
Technical support	Remained	3
Lecturer and assistant	Changed to overall satisfaction and other supports	5
Supporting function	Category deleted	0

Even though the purpose of the modification was to expedite the evaluation process, it became difficult to evaluate some important items which would contribute to the improvement of the e-learning programs, including on-line practice, synchronous discussion, effectiveness of the UI, instructor, and other support. Items dropped after some years of surveys were participation, simulator, and other support.

Contents and technical support were the two categories which were evaluated consistently over the years. Close examination of the questionnaires and the results of the interviews provided valuable insights for the improvement of the components of e-learning program. As K university started to produce and use e-learning contents most of the dissatisfaction was out of the misspelling and other grammatical errors which have been found and corrected without much delay. However, even the grammatical errors have been corrected the scores for the contents remained around the same score as before, showing as high as 4.28 in the year of 2006 and as low as 4.00 in 2008(not counting the result of the first year). One reason of the steady scores for the contents category is that the university develops new e-learning programs every year, thus making the students find some flaws with the contents every year. In addition to that reason, the number of learning units and length of them are turned out to be the major sources for the dissatisfaction with the contents. It was also related to lack of involvement with the contents activities and lack of proper prior knowledge. Whereas scores of contents seemed to be steady, the scores of technical support have been improved over the years, reaching over 4.3 from 2007. This may due to the fact that support personnel were trained to follow all standardized process of the operating services and were able to apply new support strategies.

The categories whose number of the question had been decreased greatly seem to have validity problem (Gredler, 1996). As the number of the questions was decreased to one, it's not possible to evaluate the various aspects of the category sufficiently. This study found some of the features of the original reaction evaluation have been changed as the questionnaires were modified since 2006 as shown in Table 5.

Table 5. Comparison of the features of the questionnaires

Characteristics	*Original questionnaire*	*Simplified questionnaire*
Nature of question	Detailed and investigative	General
Topics covered	All aspects	Major aspects
Purpose of evaluation	Improvement	Administrative reporting
Intended use of the findings	Apparent and informative	Need further examination

Conclusion

When we design reaction evaluation, as with other level evaluation purposes of the evaluation and the intended uses of the findings from the evaluation should be manifested (Sanders, 1994). To improve e-learning programs and system, we need to evaluate all aspects of it, such as infrastructure, contents, implementing strategy, and supporting functions. Reaction evaluation can be a good tool to measure the perceptions of the students and trainees about what they had experienced during the e-learning process which requires evaluating most of the components of e-learning

programs.

In the early years K university the categories of the questionnaire were various enough to cover most of the components of the e-learning programs. However as it repeated the reaction evaluation annually, the numbers of the categories and questions were decreased. In this case the evaluator should collect all relevant information which can be acquired from the important questions of e-learning programs. One of common errors of the evaluators is to collect information due to the convenience rather than necessity of the items.

To identify the purposes of the reaction evaluation and incorporate them to the evaluation questionnaires the evaluators should establish evaluation plan and share with the client and other major stakeholders (Patton, 1997). Through the meetings with them the scope of the information to be collected can be determined. Then the evaluator should ensure effectiveness of each stage of the reaction evaluation which include development of the evaluation plan, design of the evaluation, collection of the information, analysis of the results, interpretation of the findings, and report the results (Kirkpatrick, 2006). Although there are some matter of trade-offs between comprehensiveness and selectivity among the topics the evaluator should collect unprejudiced and reliable information.

References

01. Dixon, N. M. (1996). New routes to evaluation. Training & Development Quarterly, 50(5), 82-85.
02. Gredler, M. E. (1996). Program Evaluation. Englewood Cliffs, NJ: Prentice-Hall.
03. Kirkpatrick, D. L. (2006). Evaluating Training Programs: The four levels. San Francisco, CA: Barrett-Koehler.
04. Lee, S. H., & Pershing, J. A. (2002). Dimensions and design criteria for developing training reaction evaluations. Human Resources Development International, 5(2), 175-197.
05. Patton, M. Q. (1997). Utilization-Focused Evaluation: The new century text. Thousand Oaks, CA: Sage.
06. Sanders, J. R.(Ed.). (1994). The Program Evaluation Standards: How to Assess Evaluations of Educational Programs (2nd ed.). Thousand Oaks, CA: Sage.
07. Sugrue, B., & Kim, K. (2004). American Society for Training & Development's 2004 state of the industry report. Alexandria, VA: American Society for Training & Development.
08. Tannenbaum, S. I., & Yukl, G. (1992). Training and development in work organizations. In M. R. Rosenberg & L.W. Porter (Eds.), Annual Review of Psychology. Palo Alto, CA: Annual Reviews.