

Evaluating E-learning Programs: An Adaptation of Kirkpatrick's Model to Accommodate E-learning Environments

Thair M. Hamtini

Computer Information Systems Department, Jordan University, Amman 11942, Jordan

Abstract: Problem Statement: Kirkpatrick's model for the evaluation of training programs has been a staple in institutional learning since 1959 and is easily applicable to any training program. **Approach:** This model, however, was developed for traditional learning environments and has been regarded as antiquated, especially when one took into consideration the fact that institutional learning has increasingly taken on the form of e-learning. This study proposed an adaptation of Kirkpatrick's model, which accommodated the nuances of the e-learning environment. This model proposed a tri-stage mode of evaluation. **Results:** The three stages were interaction, learning and results. The interaction stage took into consideration the special challenges posed by the environment while the learning and results stages examined the alignment between the curriculum and the needs of an organization. **Conclusions/Recommendations:** The research conducted supported the thesis that existing training models fail to accommodate for e-learning environments and, in establishing important guidelines and criteria for the remediation as such, addressed the initial concern. The proposed evaluation method is one that is rudimentary in nature and holds a great promise for practical application.

Key words: E-learning, e-training evaluation, kirkpatrick's model, e-learning effectiveness

INTRODUCTION

The use of online technology in the learning context is still in its infancy, but it is evolving rapidly. Employees, whether at home or in the office, can access training at the exact time it is needed. This just in time training facilitates continuous access to the most current data and allows individuals more control over their learning process^[12]. Some of the other benefits of e-learning include:

- Lowered cost (cuts travel expenses, reduces the time it takes to train people and eliminates or reduces need for classroom training)
- Enhanced business responsiveness (e-learning can reach an unlimited number of people simultaneously)
- Consistent messages (everyone gets the same content, presented the same way)
- Timely content (e-learning can be updated instantaneously, making the information more accurate and useful for a longer period of time)
- 24/7 learning (people can access e-learning anytime, anywhere)

While some companies are satisfied with their e-learning programs, most online learning has been disappointing not only to learners, but also to those who

built the program. According to Rosenberg^[11], some of the reasons include the content being poorly developed. Sometimes, the e-learning content was just plain incorrect, inappropriate for the audience and purpose, or just out of date...the generic nature of the program left learners with huge holes in their knowledge. Additionally, the rush to online learning has produced some great looking yet awful training and the learning was not reinforced. Reinforcement is one of the most powerful tools for learning, as well as one of basic tenets of andragogy^[11]. Some ways to reinforce learning are to apply what you've learned on the job or having follow-up training. Lastly, the training was just plain boring. Often, the material wasn't very interesting due to layout of the e-course, text that was not relevant or exercises that didn't challenge e-learners.

Additionally, despite the instant availability of training, there is a high drop out rate. Certain individuals lack the motivation and will-power to succeed in a self-study program and high drop-out rates have been a feature of e-learning programs^[6]. While students have the option to study at their desks or at home, interruptions from colleagues and the telephone may be a problem in the office. In a survey undertaken by Training Magazine in October 2000, the main issue preventing effective delivery of e-learning was interruptions at the desktop. Others

Corresponding Author: Thair M. Hamtini, Department of Computer Information Systems, Jordan University, Amman 11942, Jordan. Phone No. 00962 6 5355000 Ext. 22611

responded that there was lack of time. Some said they could access the courses only through the company's intranet, so they couldn't finish their assignments from home. Additional issues mentioned were:

- Lack of management oversight
- Lack of motivation
- Problems with technology
- Lack of student support
- Individual learning preferences
- Poorly designed courses
- Substandard/inexperienced instructors

Today, more and more training departments are being held accountable for their work and how it affects the company as a whole. Managers and CEO's especially want to know what benefits they can expect to gain from evaluation of training. In order to examine the evaluation of e-learning, it is prudent to first operationally define evaluation and then subsequently expand the scope of this body of research from there. The American Society of Training and Development (ASTD) offers one of the clearest and most succinct operational definitions of evaluation. According to the ASTD evaluation is any systematic method for gathering information about the impact and effectiveness of a learning offering. Results of the measurements can be used to improve the offering, determine whether the learning objectives have been achieved and assess the value of the offering to the organization^[1].

HISTORY

Evaluating institutional learning:

Kirkpatrick's model: In 1987, Bell and Kerr found that out of 286 surveyed companies, less than 12% evaluated their management training programs^[2]. Another study conducted by the American Society of Training and Development found that while 90% of trainers believe in the value of evaluations, an astounding 90% of those who believe in evaluations do not use it themselves. The sole reason for not conducting an evaluation of the training was due to the fact that their companies did not require them to do so^[1]. This begs the question as to exactly what constitutes an evaluation of both traditional and electronic based training. Mager^[5] defines evaluation as the act of comparing a measurement with a standard and passing judgment on the comparison. Furthermore, he defines the differences between measurement and evaluation. According to Mager^[5], measurement is the process that determines the extent of some characteristics associated with an object or person. Mager^[5] gives the example of determining the weight of an object as measuring.

According to most training experts, evaluation is a systematic process to determine the worth, value, or meaning of something. The question of what to evaluate is

crucial to the evaluation strategy. All of this depends on the type of Human Resources Development program, the organization and the purposes of the evaluation. The information collected and used for evaluation can usually be grouped into different categories. Some methods of evaluation are more appropriate for the different categories. One founding model of evaluation is that of Donald Kirkpatrick^[4]. Kirkpatrick's model is the most well-known and widely used framework for classifying areas of evaluation. In his model, he developed a conceptual framework to aid in determining what data are to be collected. Kirkpatrick's model calls for 4 levels of evaluation and gives answers to very important questions. The four levels of this model are as follows:

- Reaction
- Learning
- Behavior
- Results.

At each level, it is prudent for the evaluator to ask certain questions. On the reaction level, it is prudent for the evaluator to ask if the participants were pleased with the program and to assess the correlates of this pleasure. On the learning level, an examination of the content of the material learned is most desirable. On the behavior level, it is important to assess whether the information learned impacted a behavioral change in the learner and on the results level, it is prudent to question whether the changes impacted as a result of training proved to be beneficial or detrimental to the organization.

Kirkpatrick defines "reaction" as what the participants thought of the particular program, including materials, instructors, facilities, methodology and content^[4]. This evaluation does not include a measure of the learning that takes place. Responses to reaction questionnaires help to ensure against decisions based on the comments of a few very satisfied or disgruntled participants. Most trainers believe that initial receptivity provides a good atmosphere for learning material in the program but does not necessarily lead to high levels of learning. "Learning" is concerned with measuring the knowledge principles, facts, techniques and skills presented in a program. It's more difficult than measuring reaction^[4]. Further, the measures must be objective and quantifiable indicators of how the participants understood and absorbed the material. They are not necessarily measures of performance on the job. "Behavior" is used in reference to the measurement of job performance. Just as favorable reaction does not necessarily mean that learning will occur, superior achievement in a program does not always result in improved behavior on the job^[4]. There are many factors, other than the training program, that can affect on-the-job performance. Lastly, evaluations at the results level are used to relate the findings of the program to organizational improvement.

Some of the results that can be examined include cost savings, work output, improvement and quality changes^[4].

Treadway Parker's model: Another way of classifying types of evaluation according to the information collected comes from Treadway Parker. Along the line of Kirkpatrick's model, Parker has divided the information evaluation studies into four (4) groups:

- Job performance
- Group performance
- Participant satisfaction
- Participant knowledge gained

According to Parker, "job performance" evaluates the extent to which an individual improved on the job. More specifically, it determines to what extent an HRD program has contributed to this improved job performance^[8]. Evaluation can come from objective measurements of job performance, including work output, quality, timeliness and cost savings. Additionally, observable changes in on-the-job behavior could be an indication of improved job performance. "Group performance" is a type of evaluation that determines the impact of a training program on a group within which the participants function, or possibly the effect of the program on the entire organization as a whole^[8]. Group performance is a difficult area to evaluate due to the many factors besides training that can affect the performance of the work group. Types of evaluation data include group performance measures of overall productivity, such as output, error rates, costs, absenteeism, and similar data. "Participant satisfaction" is a type of evaluation that determines how pleased the participants are with the training program. According to Parker, the satisfaction covers the content of the learning program, methods of training and their attitude toward what has been learned. Lastly, "participant knowledge gained" is a type of evaluation that determines what facts, techniques or skills were absorbed by the participant. In this type of evaluation, a pre-and post-training knowledge test is sometimes appropriate to measure the knowledge gained. Parker further explained that if a particular skill is to be learned, skill practices or simulations are useful for the participants to show what has been acquired^[8]. Further, according to Parker, most evaluation studies concentrate on the last two categories: participant satisfaction and participant knowledge gained. Much less frequently do they fall into the categories of job or group performance.

Jackson and Kulp's model: A slightly different approach was developed as a result of a study at AT&T and the Bell System units. Stephanie Jackson and Mary Jo Kulp^[3] presented their classification of results in an ASTD conference on determining the Payoff of Management Training. The following levels of program results or outcomes were presented:

- Reaction outcomes
- Capability outcomes
- Application outcomes
- Worth outcomes

First, reaction outcomes presents the participants' opinions of the program as a whole or as specific components such as program content, materials, methods, or activities. In summary, did they accept the program? Next, "capability outcomes" covers what participants are expected to know, think, do, or produce by the end of the program. "Application outcomes" involves what participants know, think, do, or produce in the real-world setting(s) for which the training program has prepared them. Finally, worth outcomes are the most significant result because it shows the value of training in relation to its cost. This outcome represents the extent to which an organization benefits from training in terms of money, time, effort, or resources invested.

Bridging the gap between evaluation models: Overall, the common link among the evaluation theorists and the most important element in any framework, is the ultimate outcome, which results from improved group performance. There are several ways in which to evaluate training. These evaluation instruments include:

- Questionnaires
- Attitude surveys
- Tests
- Interviews
- Observations
- Performance records

An evaluation instrument is a data-gathering device administered at the appropriate stages in the training process. Yet, whatever the instrument is, it must be reliable and valid^[7]. The questionnaire is the most common form of program evaluation instrument. The survey can be used to obtain subjective information about participants' feelings as well as to document measurable results. There are five different types of questionnaires^[7]:

- Open-ended (unlimited answers)
- Checklist (participant checks those that apply)
- Two-way questions (alternate responses)
- Multiple-choice question (several choices in which participant selects one)
- Ranking scales (participant ranks a list of items)

Attitude surveys represent a specific type of questionnaire with several applications for measuring the results of a training program. Before and after program measurements are required to show changes in attitude.

Further, measuring attitudes is a complex task. Attitudes may change over short intervals and the attitudes expressed may not always represent the participant's true feelings. In addition, the behavior, beliefs and feelings of an individual will not always correlate^[7]. Understanding that there are shortcomings, it is nevertheless possible to get a reasonable assessment on the attitude of the individual. Surveys as well as interviews and observations are good ways to measure attitudes.

There are several types of tests used in the training field. These include essay tests, objective tests, norm-referenced tests, criterion-referenced tests and performance tests. The last three types of tests described are the most common forms used in the field of training. Norm-referenced tests compare participants with each other or to other groups rather than to specific instructional objectives. They are characterized by using data to compare the participants to the norm^[9]. A criterion-referenced test is an objective test with a predetermined cut off score. Criterion-referenced tests assess whether or not participants meet the desired minimum standards, not how participants rank in reference to some group's performance. Next, performance testing allows the participants to exhibit a skill that has been learned in a training program. This skill can be verbal, analytical, manual, or a combination of the three. Performance testing is used most frequently in job-related training where the participants are allowed to demonstrate what they have learned^[9].

MATERIALS AND METHODS

Evaluating e-learning-a new model: After having delineated some of the applicable theory with regards to the evaluation of organizational learning, it is prudent to state that the lion's share of the aforementioned methods were devised for traditional learning environments and do not readily apply to e-learning environments. In order to adapt these methods for e-learning it is prudent that these methods be modified to suit the demands of an e-learning environment. In this vein, the researcher proposes an adaptation of Kirkpatrick's model. In so doing, the proposed model would be one that contains three clear, concise and self-contained focal areas:

- Interaction,
- Learning
- Results.

For each of these focal areas, the materials used will, of course, be determined by the nature and role of the area concerned (interaction, learning, results). As such, the researcher will rely upon prior research conducted in the

field, literature available from various e-learning facilities, and firsthand observations as to the nature of this e-learning environment.

Additionally, and perhaps most vitally, students of e-learning institutions will provide material in the form of valuable insight. This is constituted by their survey responses, reflections, and the researcher's analyses of their learning curve throughout the e-learning experience.

In order to ensure that this study possesses ecological validity, the methods, materials, and setting of the study will approximate a real life scenario in its reliance upon firsthand evaluation and observation, as well as literature that is focused on the issues discussed.

As with the materials, methods will be used in conjunction with the three focal areas previously established. The interaction phase would be one that gauges the ease of utility of the e-learning interface, its aesthetic qualities, user satisfaction and interaction as well as the ease at which the interface facilitated learning. This evaluation would be conducted utilizing Likert scale survey questions as well as open-ended questions. The main goal of this phase is to determine the efficacy of the translation from traditional modes of learning to an e-learning environment and to assure that the learner is afforded each and every opportunity to learn the content irrespective of advanced computer knowledge or skills. Essentially, the e-learning environment should be one that is simple to navigate and rudimentary in nature.

The learning phase is one that would measure the actual learning that occurs as a direct result of the e-course. During this phase, it is prudent for the evaluator to assess whether the learner has learned the information or has acquired the skills necessary to excel in the pertinent area. Learning can be assessed by utilizing pre and post tests. A pre-test can be administered before the actual e-learning module and a post-test can be administered within a predetermined time after the administration of the learning module. In so doing, one can determine if significant learning has occurred. This has to be closely aligned with the goals of the organization.

RESULTS AND DISCUSSION

The final stage of this evaluation model is that of results. The results phase of this model is one that has been modeled after the results phase of Kirkpatrick's model. In so doing, it examines the relative cost/benefit of the knowledge acquired, the ability for an employee to function effectively and efficiently after the prescribed training as well as the overall intrinsic and extrinsic benefits for both the employee and employer.

Three stages of e-learning evaluation based on an adaptation of Kirkpatrick's four levels of training evaluation

Stages	Type of evaluation	Characteristics	Evaluation methods
1	Interaction	Interaction can be operationally defined as the ease at which the user is able to manipulate the learning environment and to achieve the learning outcomes. It examines whether: The user was able to utilize the interface in order to learn the necessary information. In so doing, it addresses questions such as: Were the trainees able to learn the material in a fairly straight-forward manner? Was the technology conducive to learning? Was the learning environment enjoyable? Was the amount of effort needed to learn the material fair? Can the material learned in the course be extended to the work environment or be useful in other venues?	Surveys or reactive questionnaires Feedback forms Verbal or written reports
2	Learning	Learning is operationally defined as the increase in knowledge as a direct result of having engaged in the e-learning activity. In an attempt to assess whether learning has occurred, it is prudent to ask questions such as: Did the individual students learn the material they should have learned after the e-learning module? Could the theoretical knowledge attained be practically applied? What is the net change in knowledge as a direct result of having taken the course?	Pre and post course evaluations Direct observation Comparative analysis of a group of students within the e-learning environment Comparative analysis of students in traditional learning environments with students in e-learning environments
3	Results	Results evaluation can be seen as an examination of the outcome of the e-learning venture. In so doing, it examines the efficacy of the e-learning module the ability of the student to practically apply the theoretical knowledge acquired. In assessing this, it is prudent to ask questions such as: Were the skills taught easily transferable to an employment setting? Was there a marked change in the way in which the employee functioned on the job? Was there an underlying cost/time saving by the organization as a direct result of the knowledge acquired by the trainee?	Performance appraisals Direct observation by management which can be tied into pre-existing short-term employee efficiency programs Performance analysis of employee productivity

Findings in these regards reveal a confirmation of the thesis upon which this work was based: Kirkpatrick's model, like many others of its nature and stature, is not readily applicable to e-learning without a careful reevaluation of the model. Such a reevaluation, considering the materials available in relevant literature on the subject and firsthand interactions with e-learning students, reveals that the overall satisfaction and efficacy of e-learning still lacks what in-person training can offer.

Kirkpatrick's evaluation strategies represent one of the most comprehensive strategies for evaluating organizational training. It is one that operates on the presumption that the return on investment is one of the most fundamental aims of training initiatives. In this vein, there have been numerous studies which indicate that e-learning programs are as effective as their traditional counterparts when one examines the learning outcomes. More specifically, these studies indicate that there is no significant difference in learning outcomes (Level 2 of Kirkpatrick's model) when compared to traditional classroom instruction. However, when one examines trainee satisfaction (at level 1), one can clearly see that the recipients of face-to-face instruction have expressed more satisfaction^[10]. Much of this can be attributed to the interface and the impersonal nature of the e-learning environment. These have been

significant challenges when one attempts to extend training to the e-learning environment.

CONCLUSION

Overall, there are several clear and concise methods for evaluating learning both traditional and e-learning. These methods have been designed for traditional learning environment but have been adapted for e-learning. There adaptation was problematic at times in that the lion's share of the methods are theoretical in nature but their practical application is one that has been illusive at times. The proposed evaluation method is one that is rudimentary in nature and holds a great promise for practical application. It is one that takes into account the strong influence of the interface as well as the necessity to adapt training methods to the needs of the human resources departments within organizations. In so doing, it assures that learning does take place within the organization and that learning is of mutual benefit to both the organization and the individual employee.

This is, admittedly, no matter what evaluation module is concerned and what institution is in question, the relevant issue: in producing learned, satisfied students the organization in question will ultimately profit from the transaction, be it in-person or via e-learning.

REFERENCES

1. American Society of Training and Development (ASTD), 2006. ASTD's Annual Review of Trends in Workplace Learning and Performance. http://www.astd.org/NR/rdonlyres/0A1BE935-3905-4B09-B517-6CC5B41E2AC5/12315/stateofindustry_memberonly.pdf.
2. Bell, J.D. and D.L. Kerr, 1987. Measuring training results: Key to managerial commitment. *Train. Develop. J.*, 41: 70-73. http://eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_ERICExtSearch_SearchValue_0=EJ344830&ERICExtSearch_SearchType_0=no&accno=EJ344830
3. Jackson, S. and M.J. Kulp, 1979. Designing Guidelines for Evaluating the Outcomes of Management Training. In Peterson, R. O. (Ed.), *Determining the payoff of management training*. Madison, WI: ASTD, ISBN: 10:9992589248.
4. Kirkpatrick, D. and Kirkpatrick, P. 2006. *Evaluating Training Programs*. 3rd Edn, Berrett-Koehler Publishers, San Francisco, CA., ISBN: 10: 1576753484.
5. Mager, R.F., 1984. *Measuring Instructional Results*. 2nd Edn., David S. Lake Publishers, Belmont, CA., ISBN: 10: 0822443406.
6. McDermott, T., 2001. Corporate training: Traditional vs. online. *Account. Ireland*, 33: 30-31.
7. Morris, L.L. and C.T. Fitz-Gibbon, 1978. *How to Measure Program Implementation?* 1st Edn., Sage Publications, Beverly Hills, CA., ISBN: 10: 0803910665.
8. Parker, T.C., 1973. Evaluation: The forgotten finale of training. *Personnel*, 7: 61-63. ISSN: 0031-5745.
9. Payne, S.L., 1951. *The Art of Asking Questions*. Princeton University Press, Princeton, NJ., ISBN: 10: 0691093040
10. Phillips, J., P. Phillips and L. Zuniga, 2000. *Evaluating the Effectiveness and the Return on Investment of e-Learning*. American Society of Training and Development, Alexandria, Virginia. <http://www.neiu.edu/~sdundis/hrd490/evalqstns.pdf>.
11. Rosenberg, M.J., 2001. *E-Learning: Building Successful Online Learning in Your Organization*. 1st Edn., McGraw Hill, New York, USA., ISBN: 0071362681, pp:344.
12. Sadler-Smith, E., 2001. A reply to Reynold's critique of learning style. *Manage. Learn.*, 32: 291-304, DOI: 10.1177/1350507601323001.