

* *Ridha M. Al-Khayyat*

** *Abdulaziz A. Taqi*

* Institute of Banking
Studies, Kuwait

** The Public Authority
for Applied Education,
Kuwait

TWO STEPS TRAINING AND DEVELOPMENT REACTION EVALUATION

Key Words

*Training and
Development;
Training Reaction
Evaluation;
Two Steps Training
Reaction Evaluation;
Four Level Training
Evaluation.*

Abstract

Training reaction evaluation is the most used level of Kirkpatrick's four levels training evaluation model. In order to better understand the factors influencing adult trainees' reaction, a two step reaction evaluation model was developed and empirically tested. The daily evaluation (i.e., first step) variables, the time effect variable, and two sets of dummy variables reflecting different programs classifications were found to be related to the final evaluation (i.e., second step) variables. This leads to the conclusion that trainees' reaction is not based on the events of the last day of the training program, rather it is a result of trainee exposure and reaction of the daily events of the training program. Regression analysis further substantiated this finding. The daily program evaluation outcomes were found to be a good predictor of the trainees' eventual assessment of the training experience as a whole, i.e. trainee's reaction. The final program evaluation factors, with the exception of the program's administrative arrangement, could also strongly explain the changes in the final overall trainees' reaction.

Introduction

There is unanimous agreement in the literature, supporting the view, that evaluation of training is essential. What remains is the ques-

tion which type of evaluation method is most appropriate to measure training effectiveness? It is not the intention of this Paper to present an analysis of

the various approaches available in the literature, however, it is the intention to explore in detail one recognized approach. That is reaction evaluation, or first level of evaluation (Kirkpatrick 1959a, 1959b, 1960a, 1960b, 1976, 1996). It is an acceptable method of evaluation, that determines the level of participants' satisfaction (and/or positive reaction), and therefore the extent to which participants have learned the skills, knowledge and attitudes which it has sought to impart. Inherent in this approach is a relatively acceptable assumption that the more positive the reaction of the participant the more likely the occurrence of learning. The question then arises of how to achieve higher trainee positive reaction and/or satisfaction.

This Paper attempts to study the dynamics of trainees' reactions, the factors that explain these reactions, and the relationship between these factors. The aim is to better understand trainees' reaction and therefore understand how to achieve higher positive reactions. A model is developed in order to achieve better understanding of trainees' reactions.

Research Objectives

The main objectives of this study are: (1) to develop a model explaining variables forming and/or influencing trainee's evaluation and/or reaction to a training program; (2) to empirically test

the relationship between the variables included in the proposed model in an attempt to better understand trainee evaluation and/or reaction to a training program.

In the remaining portion of this Paper a reaction evaluation model is presented, empirically tested and results reported.

Evaluation of Training and Development Activities

It might be said that, although every training manager wants to measure exactly where he/she is heading (i.e., evaluate training), training and development evaluation is not the most attractive option. Several authors agree that evaluating training is not wanted, because it triggers the fear of those responsible for implementing training activities (See, for example, Tracey 1981; Goldstein 1974; and Kirkpatrick 1983). The fear of evaluation coupled with the high cost of training evaluation and its complexity (Dionne 1996), therefore, diminishes the desire to pursue training evaluation.

Kirkpatrick's four levels of training evaluation helped to diffuse some of the difficulty. It is widely used in the literature and among practitioners, for it provides a simple framework of how to think about evaluation and how to perform it. The four levels are: reaction, which is the measurement of how parti-

participants feel about the training program; learning, which is the function of determining the extent to which participants acquired the skills, knowledge and attitudes targeted by the training; behavior, which is concerned with finding out whether the trainees' level of work performance had changed due to the training program; and results, which is the quest of determining the final results in terms of organizational productivity.

The popularity of this model is such that, the literature either revolves around it, or in it. The majority of the literature tends to revolve in it, seeking to either build on it or validate it (Hamblin, 1974; Kaufman and Keller, 1994; Phillips, 1995; Brinkerhoff, 1987; Faerman and Ban 1993). The minority tends to revolve around it by questioning it (Alliger and Janak 1989) and some times reject it and calls for a more solid theory of training evaluation (Holton 1996, Dionne 1996).

According to Alliger and Janak (1989), three implicit assumptions about this model seem to appear in the minds of researchers and trainers. First, the so-called levels or steps are arranged in ascending value of information provided. This is to say that a measure of learning provides more information than does a measure of reaction, and so forth. Second, the levels or evaluation are causally linked. Hamblin (1974) contends training leads to reactions, which lead to learning, which leads to changes in

job behavior, which leads to changes in the organization. Finally it is assumed that the levels are positively inter-correlated. Alliger and Janak (1989) further concluded that these three assumptions can be logically questioned.

Furthermore, it seems that, because of the second and the third assumptions, practitioners are more frequently using the reaction evaluation (i.e., the first level). However there exists a strong debate between researchers in the literature relating to the importance of reaction evaluation. Some believe that it is important and a prerequisite to any evaluation activity (Kirkpatrick, 1996; Faerman and Ban, 1993), while others believe it has an indirect relation to learning (Holton 1996; Alliger, and Janak 1989) and therefore more focus should be given to the other three levels of evaluation.

Trainee Reaction Evaluation

The majority of the studies in training evaluation literature attempt to study trainee reaction in terms of its relation and/or impact on other factors such as learning, behavioral changes, and organizational results (see for example, Holton, 1996; Mathieu, Tannenbaum, and Salas 1992; Alliger and Janak 1989; Faerman and Ban 1993). That is to say that these studies attempt to validate and/or critique Kirkpatrick's four levels of training evaluation. There are few

studies that articulate what constitute training reaction, and they are usually found in the education rather than training related literature (see for example Mahmoud, 1991). Of these studies one attempting to manipulate trainees' choices concerning whether to attend a training program, found that those given the choice demonstrated more positive reactions to the program (Hicks and Klimoski, 1987).

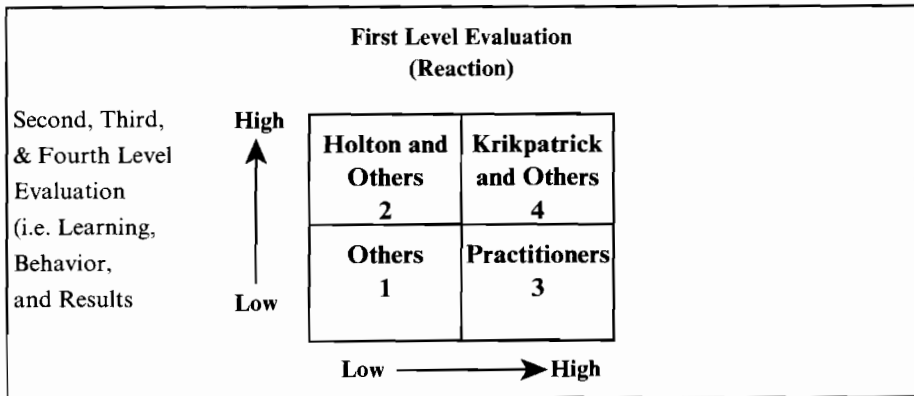
The literature on reaction evaluation can at best be described as strewed. In the following, I will utilize the quadrant analysis technique in order to synthesize and better understand the current literature.

As depicted in Figure (1), the view in the first quadrant is that neither reaction evaluation nor the other three evaluation levels (i.e., learning, behavior, and results) are important. That is to view, evaluation of training as unnecessary, since its bene-

ficial effects are obvious. It is threatening for it may reveal inadequacies, and it is expensive especially when conducting a scientific evaluation with cause and effects determinations (Truelove, 1995).

The second quadrant, where reaction evaluation is low and the other three levels of evaluation is high, represents the view that trainee reaction is unrelated to learning (Patrick, 1992) and that reaction, is at best a function of the relationship between training motivation and learning (Mathieu, Tannenbaum, and Salas 1992). Therefore reaction is an aid to learning (Holton, 1996). According to this view the focus should therefore be given to the other three evaluation levels, where the benefit of training is measured in terms of its impact on learning, behavior, performance and organizational results.

Figure 1
The Importance of Reaction Evaluation



The third quadrant, where reaction evaluation is high and the other levels of evaluation is low, represents the practitioner's perception of training evaluation and the importance of reaction type evaluation. The view is based on the premise that since rigorous scientific evaluation of training, where its impact is quantitatively measured is expensive and requires the allocation of various resources, reaction evaluation is sufficient enough to guide training managers in making decisions on how to improve the training and justify it as a future investment. After all, the benefit of training is so obvious. One has to note that few organizations undertake training activities without some sort of evaluation exercise to determine the effectiveness of the training. The advantages of reaction evaluation are (1) it requires very little time and effort, (2) it is tolerated by the client, (3) it is relatively unobtrusive, and (4) it provides participants with a sense of contribution into subsequent training (Tannenbaum and Woods 1992). Although several authors argue about the usefulness of this type of evaluation, it is the most used form of training evaluation (Goldstein, 1980; Saari *et al*, 1988; Swierczek and Carmichael, 1985).

Finally, the fourth quadrant, where reaction evaluation is high and the other three levels of evaluation are

high, depicts Kirkpatrick level evaluation (1959a, 1959b, 1960a, 1960b, 1976, 1996) and those who made some modifications to it (Hamblin, 1974; Kaufman and Keller, 1994; Phillips, 1995, and Brinkerhoff, 1987). Here reaction evaluation is seen as an integrated element of the training function. Reaction evaluation is given the highest importance compared to the other quadrants. It is a prerequisite for determining that the effectiveness of training, has a causal link with the other three levels, and is assumed to be inter-correlated with the other levels (Hamblin, 1974; Faerman and Ban 1993; Clement, 1982; Noe and Schmitt, 1986; Wexley and Baldwin, 1986). The view in this quadrant is that reaction is related to learning. Kirkpatrick contends that if training is going to be effective it is important that trainees react favorably. It is therefore necessary to achieve higher levels of favorable trainee reactions in order to achieve favorable outcomes at the next remaining three levels (Clement, 1982).

In the final analysis, it is clear that reaction evaluation is highly recommended in the fourth quadrant, extensively used by practitioners who belong to the third quadrant, and considered as an aid to learning i.e., second quadrant. Even for those who might believe that evaluation of training is not necessary (i.e., first quad-

rant) they tend not to mind using the reaction evaluation. It is based on this conclusion that this paper attempts to better understand what factors influence and construct trainees' reaction to training activities. This study, however, is not about whether trainee reaction impacts and/or correlates with learning, behavior, and results (i.e., the other three evaluation levels), but only about what factors lead to the formation of trainee reaction during the program. The findings will help practitioners to better design a learning environment that leads to higher trainee satisfaction. It should also provide a thorough understanding of the nature and dynamics of trainee reaction, fill a gap in the training and development literature, and aid future research attempting to measure the relationship between reaction and learning.

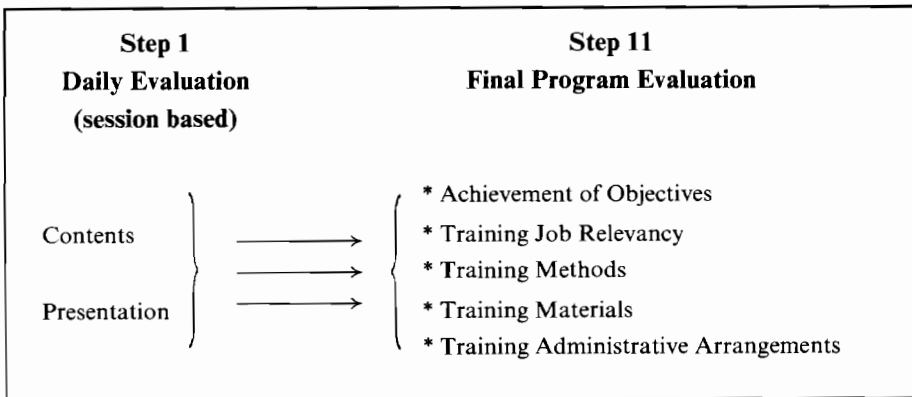
Two Steps Reaction Evaluation Model

Figure (2) depicts the suggested Model that attempts to explain various factors leading to the formation of participant's reaction of the training event. Reaction evaluation is performed in two steps. First, trainee's reaction is measured on a daily basis and for each session of the training program. In this step, two factors are assumed to impact participant's reaction and lead to the formation of the overall program reaction:

- (1) session's content and (2) session's presentation.

Second, in the final day of the program, trainee's overall program reaction is measured. Five factors are assumed to form the participant's reaction; (1) achievement of objectives, (2) training job relevancy, (3)

Figure 2
Two Steps Training Reaction Evaluation Model



training methods, (4) training materials, and (5) training administrative arrangements. It is assumed that these five factors are accumulation of participant's daily reactions towards the presentation and the contents. That is to say that participant's overall program reaction is not an isolated event, formed at the end of the program, but an accumulation of the program daily reactions.

It should be noted that, the factors in this model are found in various training and development models in the literature (see for example Al-Khayyat and Elgamal, 1997; Nadler, 1983; Patrick, 1992; Camp *et al.*, 1986). They are also assumed to lead to the formation of trainee's overall program reaction. A further detail of each of these factors follows.

Contents: Once the training objectives have been clearly defined, the subject matter content can be developed. According to Patrick (1992) training contents form the basis of the training program, which has to be mastered by the trainee. The nature of training contents depends on, among other things, training objective, subject matter (i.e., technical versus behavioral), time allocated for the training event and available training aids.

Presentation: In the heart of any training activity is the usage of appropriate presentation format. The latter

refers to the way messages are delivered to an audience (Rothwell and Kazanas, 1992). The presentation is usually organized in such a way that relates to the training objectives, the nature of the subject matter, and the audience educational and work experience. Instructors should at least be able to demonstrate effective presentation skills (Foshay, Silber, and Westgaard, 1988).

Achievement of Objectives: The first step, after defining the training needs of a target group is the process of setting training objectives. There is unanimous agreement in the literature to the view that without written established learning objectives, the likelihood of developing a successful training event is questionable (Camp *et al.*, 1986; Gagne, 1977; Mager, 1975). In fact there is no training model in the literature that does not include the development of learning objectives as the first step after the identification of training needs (see for example Al-Khayyat and Elgamal, 1997; Nadler 1982; Patrick, 1992; Camp *et al.*, 1986). Mager (1975: 5) presents a useful definition of instructional objectives a description of performance you want learners (trainees) to be able to exhibit before you consider them competent. An objective describes an intended result of instruction, rather than the process of instruction itself. (from Camp *et al.*, 1986: 100).

Training Relevancy: There is no doubt that adults learn better when they realize the relevancy of the techniques and/or concepts to their jobs. Realizing the relevancy means that the trainees perceived a potential application of the program's contents.

Training Methods: There are many training methods available, each with its own strength and weakness. Examples of these are lecture, case studies, role-play. The methods tend to take considerable time and effort, require different skills and require planning in order to ensure higher transfer of potential learning from the training situation to the trainee's job situation. According to Campbell *et al* (1970), training methods can be classified into three categories: (1) information presentation such as lectures and reading lists; (2) simulation methods, examples of which are case methods, role playing, and games; and (3) on-the-job training methods which includes orientation training, apprenticeships and job rotation. The choice of training methods depends on, among other things, the training objectives, contents, trainees' backgrounds and characteristics, time available and cost.

Obviously some training methods are appropriate for more than one type of training and development objective(s) or strategy. The latter is due

to the fact that there exists a difficulty in rigorously determining what behavioral objective(s) is/are likely to be modified by the various instructional techniques (Goldstein, 1986). However, generally speaking, there is no doubt that instructional strategies are more applicable for a particular training strategy and therefore will be more effective in achieving specific training objective(s). For example, the use of machine simulators are most appropriate to develop technical skills, as in flight simulators, whereas for team building, role playing, sensitivity training are more appropriate for promoting effective group interaction and other interpersonal skills.

Training Materials: This is one of the subject matters that are neglected in the literature. The majority of what is written about this subject is found in practical journals and text books, where emphasis is given to the presentation and problems with program materials (see for example Munson, 1992) rather than research applications. The training material is a hidden aspect of any training activity that is usually developed as part of the curriculum design phase of the training process. It is the hard version of the concepts and/or skills and knowledge that are expected to be transferred to the trainee. Usually the program material is found in a written format.

However, with the new technological development and the use of computer based training, program material may be found on diskettes and /or CDs, integrated with the training presentation and methods.

The trainee is expected to react to the material since it is used not only during the program but also as a reference after the program. Thus, the clarity of the material is highly important and the relevancy of the materials to the trainees' job is also of great concern.

Training Administrative Arrangements: Program administration involves a set of activities supplementing the instructional strategy and/or methods. It includes activities related to the training facilities (i.e., building and seminar halls, table and chairs) and catering services (e.g., coffee, tea, lunch etc.). The major objective of the administrative support activities is to facilitate and/or minimize obstacles, which might hinder the ade-

quate operation of the training and development activities and thus create an effective learning environment during the training program. An elaborate administrative support discussion is found in Nadler's Critical Event Model (1983).

Data and Methodology

Subjects in this study were participants in the training programs offered by the Training Unit at the Institute of Banking Studies (IBS). The programs are loosely classified under different banking related subjects. The programs were also designed for different job levels from junior to top executives. The study included data gathered from programs during four consecutive training years (1994 until 1997) with a total of 185 training programs. Table (1) summarizes the number of training programs and also the number of trainees in each training year included in the study.

Table 1
Description of Sample

Year	# of Program	# of Trainees
1994	54	658
1995	57	680
1996	58	765
1997	18	232
Total	185	2335

The reaction evaluation was performed in two steps with two different questionnaires. The first step included the daily reaction evaluation questionnaire which consisted of two Likert-type statements each of which is followed by a six-point scale that ranges from low to high. The first statement asks the trainee to evaluate the content (D_Cont) whereas the second relates to the presentation (D_Pres) of each session (See Appendix (I)).

The second step included the final overall evaluation questionnaire that consists of five sections with a total of 17 Likert-type statements and other program related questions (See Appendix II). A six-point scale followed Likert-type question statements. The 17 statements of the final evaluation questionnaire measured trainee's reaction as it relates to the five attributes of the course: achievement of course objectives - OBJ (1 statement), training relevancy - REL (6 statements), training methods - METH (5 statements), training material - MATR (3 statements), and training administrative arrangements - AMIN (2 statements). The overall daily (O_Daily) is the mean of the two daily variables, i.e., D_Cont, and D_Pres. The overall final reaction (O_All) is the weighed

mean of the five overall attributes of the training program.

The daily and final evaluations were conducted without revealing trainee's identity. For that, results for the final evaluation of each trainee could not be related to his/her daily evaluation. All analysis was then conducted on the overall results of the final and the daily evaluation. The final evaluations for each program were averaged overall trainee's evaluation. This led to one observation per program. To relate this to the daily evaluation all results of trainee's session's evaluation were averaged, and then the session's evaluations for the program were also averaged in one observation. Due to the confidentiality of the questionnaires and the fact that trainee's identity was anonymous, analysis could not be performed on gender, education background, years of experiences, etc. In addition such factors were not of the scope of the interest of this paper.

The main research objectives of this paper is to investigate whether there is correlation between the daily evaluation and the final evaluation, and if so could the daily evaluation be used to predict the final evaluation. The following statistical techniques

will be used: (1) descriptive statistics to evaluate the variables of the study; (2) Pearson's correlation coefficient to investigate the magnitude, strength, and direction of linear correlation between the daily and final evaluation variables; (3) linear regression models in which each of the five final evaluation variables is expressed by the two daily evaluations variables, the time effect variable (Year) and a set of dummy variables to reflect different program classification criterias repre-

senting nature of the program and type of trainer. All of these were used in linear models form.

Table (2) presents values of the time effect variable (Year), three sets of dummy variables (DA1, DA2, DA3) representing the nature of the program with three classifications, and another set of dummy variables (DB1, DB2, DB3) classifying the training programs according to the trainer type.

Table 2
Time and Program Classification Variables

Variable		Values
Year		94, 95, 96, 97
Dummy Variable for the Nature of the Program	DA1	1 if program aimed at increasing participant knowledge base 0 otherwise
	DA2	1 if program aimed at equipping participants with technical skills 0 otherwise
	DA3	1 if program aimed at prompting a behavioural change in the participants 0 otherwise
Dummy variable for the Type of Trainer	DB1	1 if the trainer was from outside Kuwait 0 otherwise
	DB2	1 if the trainer was from local education and/or institutions 0 otherwise
	DB3	1 if the trainer was from local banking industry in Kuwait 0 otherwise

Table 3
Reliability Analysis of Multi-Dimension Scale Variables

Scale Variable	Questions Included	Cronback's* α	Standardized Item α	Max. α If some variables delete
Training Relevancy REL	Q2, Q3a, Q3b Q3c, Q3d, Q4	0.9043	0.9050	0.9020
Training Methods METH	Q5, Q6a, Q6b Q6c, Q6d	0.8240	0.8252	0.8125
Training Material MATR	Q8, Q9, Q10	0.8434	0.8469	0.8544
Training Administration ADMIN	Q7a, Q7b	0.6933	0.6957	-

* For variables used in the final evaluation form please refer to Appendix II.

Since the scale variables REL, METH, MATR, and ADMIN were multi-dimensional formed from sets of variables included in the program final evaluation questionnaire, a reliability test was conducted to measure the consistency and reliability of variables in each set in measuring each of the above four dimensional factor. Table (3) presents the results of reliability analysis.

Based on Cronback's statistics using the original and the standardized variables, the scale variables used in forming the three dimensional factors: REL, METH, and MATR are highly to very highly consistent and reliable. The two variables forming the administrative arrangements fac-

tor (ADMIN) were acceptably reliable.

The last column in Table (3) (Maximum (α) if some of the variables are deleted) indicates that each variable in the dimensional factor is important and consistent with other variables forming the factor. Eliminating any of these variables would not improve the reliability of each of the scale variable.

Table (4) depicts the descriptive statistics for the sample data. It includes the mean, standard deviation, minimums and maximums for each variable, and percentage of satisfaction (i.e., (Mean / 6) x 100). The lowest observed mean was 4.9246 and the highest was 5.4334, which leads to the

Table 4
Descriptive Statistics of the Data

Descriptive	Minimum	Maximum	Mean	Std. Deviation	Percentage %
D_CONT	4.4949	5.9737	5.4334	.3125	90.6
D_PRES	4.1667	5.9598	5.3605	.3460	89.3
OBJ	2.6364	6.0000	5.1258	.5114	85.4
REL	2.5091	5.8792	4.9246	.5130	82.1
METH	2.8545	5.8769	4.9919	.4916	83.2
MATR	2.8519	5.8889	5.0921	.4866	84.9
ADMIN	4.0000	6.0000	5.2005	.3739	86.7
O_Daily	4.3715	5.9627	5.3970	.3207	90.0
O_All	3.1671	5.8559	5.0670	.4122	84.5

conclusion that the participants positively received the programs.

Results in Table (4) also revealed that the mean of the two daily evaluation variables mainly D_Cont and D_Pres were higher than the final evaluation variables and with smaller standard deviations. It also revealed that the percentage of satisfaction from daily sessions was about 90%, whereas the final program level of satisfaction was about 85%. This could be explained by the fact that daily variables were based on two questions in the session evaluations that should be short and overall perceptive based, whereas the final evaluation variables were five and were based on a number of more detailed

questions (variables). For future references, we will name this explanation as the questionnaire limitation effect. One additional explanation is the fact that session evaluations were focused on specific issues related to each session the trainee had an immediate experience of, whereas the final evaluation is a summary of more general issues of which trainees had less recent experience of. We will name this explanation as the trainee's memory limitation effect.

Table (5) includes the correlation coefficients between the two daily evaluation variables and the five final evaluation variables. All correlations are significant at $p = 0.01$ or better with the exception of the training

Table 5
Correlation Coefficients for the Two Daily Variable
and the Final Evaluation Variables

	D_CONT	D_PRES	OBJ	REL	METH	MATR	ADMIN	O_Daily	O_All
D_CONT	1.000								
D_PRES	.897**	1.000							
OBJ	.588**	.591**	1.000						
REL	.608**	.574**	.844**	1.000					
METH	.584**	.521**	.774**	.842**	1.000				
MATR	.569**	.531**	.796**	.857**	.830**	1.000			
ADMIN	.264**	.266**	.348**	.365**	.479**	.492**	1.000		
O_Daily	.971**	.976**	.606**	.606**	.565**	.564**	.272**	1.000	
O_All	.619**	.587**	.894**	.928**	.923**	.934**	.589**	.618**	1.000

** Correlation is significant at the 0.01 level or less).

administration variable. The positive correlation coefficient indicates that the change in one variable implies a similar change in the direction of the other variable. As far as the value and magnitude of the correlation is concerned, we will consider that values above or equal to 0.5 indicate relatively moderate to strong linear relation. There are several high correlations, take for example, the relation between D_Pres (.897) and D_Cont suggests that the better the trainer's presentation skills, the richer the content he/she could deliver during sessions. This is also substantiated in the relation between D_Cont and O_Daily (.971) and also the relation (.976) between D_Pres and O_Daily

(.976). The high correlations between achievement of objectives and training relevancy (.844), training methods (.774), and training material (.796) means that the latter three variables are strongly related to whether the program will achieve its objectives.

One observation of interest is the relatively high correlation (.618) between O_Daily and O_All. This means that there exists a relatively strong relation, between the daily reaction of the trainees with the presentation and content on one hand, and overall mean of the five variables of the final evaluation the higher the former, the higher the latter. This is to say that the trainees' overall reaction to the program is not an isolated event that is

developed and perceptually formed on the last day of the training program. It is a result of an accumulated daily perception of the trainee that leads to the final overall perception and/or reaction. One can then refute the notion that trainees tend to positively and quickly mark the items in the final questionnaire, either because they feel happy to have finished the program, or rushing to leave in order to cater for other pressing personal needs.

One surprising observation is the impact of the training administrative arrangements in forming trainees' reaction. The low correlation of the training administration with the D_Cont (.264), D_Pres (.266), achievement of objectives (.348), training relevancy (.365), methods (.479), and materials (.492), means that this variable although has a moderate to strong relationship (.589) with the O_All (i.e., overall five variables), it does not have a strong relationship with each of these variables taken individually.

Having in mind the fact that in the particular area researched the catering services for these programs was provided by a five star hotel restaurant, in a sea-view location, one may conclude that the trainees tend to give great importance to the quality of the program's other attributes (i.e., contents,

presentation, relevancy, methods, and materials), rather than the catering and facilities when forming their program reactions.

Table (5) also reveals that the correlation between the two daily evaluation variables is high. A similar conclusion with the five final evaluation variables, with the exception of ADMIN, is depicted. However, the correlation coefficients between the daily valuation variables and the final evaluation variables were moderately high (between .56 and .61, with the exception of the ADMIN). This could be explained by the previously mentioned "questionnaire limitation" and trainee's "memory limitation" effects.

As previously mentioned the main objective of this study is to examine the power of the predictability of the daily evaluation variables in explaining and predicting the five dimensional factors and the final evaluation variables (OBJ, REL, METH, MATR and ADMIN). Each of the dimensional factors and the overall program evaluation variable will be regressed using the daily contents (D_Cont), the daily presentation (D_Pres), time factor (Year), the dummy variables set identifying the program nature (DA1, DA2, DA3), and the dummy variable classifying the type of trainers (DB1, DB2, DB3).

The best set of regressions command of the MINITAB version 13 was used to identify the best three models using one independent variable, the best three models using two independent variables. etc, until the model that includes all independent variables.

Among all the models presented by the above technique the best of all possible regression models are identified using the R², adjusted R², sum of squares of the residuals and the Mallows's Cp Statistics. Best models were chosen based on nearest Cp statistics to the number of parameters included in the estimated model, with the highest possible R² and adjusted R², and with the lowest sum of squared residuals.

As was indicated in Table (5) daily contents (D_Cont) is highly correlated with the daily presentation (D_Pres). To minimize the effect of multi collinearity between these two independent variables, step-wise regression was also used in estimating the best possible models with 0.10 significant level (α) for entrance and 0.15 significant level (α) for removal.

Table (6) presents the result of the regression models to explain and predict each of the overall final evaluation (reaction) factors and the aggregate overall final factor from the daily attributes (i.e., D_Cont and D_Pres).

It also includes the mean of the daily evaluation variables. As indicated previously, the study included 185 observations (i.e., programs) and 14326 daily evaluations were averaged over all participants, for all days within each program to obtain 185 observations of the two daily evaluation variables.

The Table presents the estimated coefficients for each of the daily independent variable along with time effect variable (YEAR), the set of dummy variables to identify program nature (DA1, DA2, DA3), and the set of dummy variables to classify the type of trainer in the program (DB1, DB2, DB3), in addition to the constant terms. Values within brackets indicate the calculated t-value for each of the estimated coefficients. The value of the standard error of the estimated model (SEE) is also presented for each of the seven estimated models, along with the multiple correlation coefficient (R), the coefficient of Determination (R²), adjusted coefficient of determination's (adj-R²), and the value of Mallows's Cp statistics.

Model (1) relates to the achievement of objectives (OBJ) based on the collective means of participants regarding the daily contents (D_Cont) and the daily presentation (D_Pres). Using the best set of regression command and the step-wise regression (of

Table 6
Estimated Regression Models

No.	Dependent Variable	Independent Variables				First Dummy Set			Second Dummy Set			Cp * Statistics	R ² R2adj	Mult. R	SEE	
		Constant	D_Cont	D_Pres	Year	DA1	DA2	DA3	DB1	DB2	DB3					
1	OBJ	0.110 (0.209)	0.531 (2.45)	0.414 (2.09)									3.4	0.382 0.372	0.618	0.405
2	REL	-0.0086 (-0.166)	0.962 (10.26)			-0.209 (-1.97)	-0.280 (-3.8)						3.7	0.416 0.406	0.645	0.395
3	METH	0.0047 (0.009)	0.946 (9.86)			-0.205 (-1.8)	-0.148 (-1.99)						4.2	0.383 0.370	0.619	0.390
4	MAT-R	7.059 (2.37)	0.852 (9.22)										0.3	0.369 0.359	0.608	0.390
5	ADMIN	8.677 (3.17)		0.271 (3.50)		-0.051 (-1.79)	-0.136 (-1.98)						3.3	0.103 0.083	0.321	0.358
6	O_ALL	5.730 (2.35)	0.784 (10.48)			-0.050 (-2.01)	-0.213 (-3.53)						3.1	0.427 0.414	0.654	0.315
7	O_ALL	6.188 (2.58)		O_Daily 0.770 (10.621)				0.207 (3.55)					1.0	0.424 0.408	0.651	0.317

* Cp is the Mallows Cp Statistics (By: C. L. Mallow)

$C_p = \frac{\text{Resid. SS}_p}{\text{MSE}_m} + 2(P+1) - n$; where m is the number of all independent and dummy variables

p is the number of used independent and dummy variables.

$E(C_p) = P$

the SPSS WIN version 10) the best estimated model to predict OBJ is the one presented by Model no. 1 in Table (6), in which the two daily variables, the constant term and the dummy variables representing technical skill programs are found to be significant and effective. All independent and dummy variables estimated parameters are significant at 0.05. The model indicates that OBJ is explained by 53.1% of D_Cont and 41.4% of the D_Pres, with a .137 negative contribution for the technical skill programs compared to the other programs. This means that a full point of satisfaction with content leads to .531 increase in satisfaction with OBJ, while a full point of satisfaction with presentation increases satisfaction with OBJ by .414. Further, satisfaction with OBJ is decreased by .137 for the technical skill programs. The multiple correlation coefficient (R) is 0.618. The Cp statistics was 3.4 which is less than the total number of independent variables plus the constant term (4-parameters), and this indicates that the estimated model is stable and significant. The estimated model is also not affected by multi-collinearity symptoms.

Model (2) estimates the training relevancy (REL). The best estimated model includes a constant term, the daily contents variable (D_Cont) and two dummy variables from the set of nature of the programs (DA1 and

DA2). The results clearly indicate that D_Cont explains 96.2% of the changes in REL. Knowledge base (DA1) and technical skill programs (DA2) are negatively contributing to that factor by 0.209 and 0.280 respectively. Which means that a full point of satisfaction with contents leads to .962 increase in satisfaction with REL. Further, satisfaction with REL is decreased by .209 for knowledge based programs and .280 for technical skill programs.

Model (3) depicts the training method (METH). The estimated model shows a similar set of independent variables as in model (2), namely D_Cont, DA1, and DA2 but with a new dummy variable identifying the trainers from Kuwaiti local banking industry and with 94.6% contribution from the daily content variable (D_Cont).

Model (4) relates to the training material (MATR). The best estimated model for (MATR) involves the daily contents variable (D_Cont), an effect from technical skill type of programs (DA2) and a time effect variable (Year). Daily content contribute (85.2%) to the estimated value of MATR. DA2 contribute negatively by 0.201 and time variable also contribute negatively by 0.068. This is to say that, a full point of satisfaction with contents leads to .852 increase in

satisfaction with MATR. In addition, satisfaction with REL is decreased by .201 for technical skill programs, and by .068 for time variable.

An interesting observation is found in model (5). It clearly indicates that the training administrative arrangements (ADMIN) is not sufficiently explained by the two daily independent variables (D_Cont and D_Pres) along with the time and the two sets of dummy variables. This means that we cannot sufficiently predict or explain the changes in the (ADMIN) by these variables.

In model (6), an attempt is made to estimate the impact of the set of predictor variables on the five final evaluation variables combined (i.e., O_ALL). The estimates include the effect of D_Cont, Year, DA1, and DA2. With 78.4% contribution of the daily content variable. This model indicates that a full point of satisfaction with contents leads to .784 increase in satisfaction with O_ALL.

Finally, model (7) confirms what the correlation analysis indicated. The combined D_Pres and D_Cont, that is the O_Daily, contributed 77% to O_All with positive significant contributions of .207 from programs with the aim of promoting a behavioral change in the participants (DA3), and a negative contribution by time variable (YEAR). This means that a

full point of satisfaction in O_Daily leads to .77 increase in satisfaction with O_ALL and satisfaction is also increased by .207 from programs aiming to promote behavioral change in the participants. Further, satisfaction with O_ALL is decreased by .056 from the time variable.

Table (6) also presents the multiple correlation coefficient (R) of the estimated models. These estimates reveal that (R) for each estimated model was higher than the Pearson's correlation coefficient between each daily variables and the dependent variables. Values of R were around 0.60 to 0.65 which means that there exists linear correlation between each of the dependent variables in these models and the predicting variables. However, the low values of R² for these models were because of the questionnaire limitation effect' and the trainee's memory limitation effect'.

It is worth mentioning that in all models presented in Table (6) the contribution of the technical skill programs (DA2) and the knowledge-based programs (DA1) were negative (even though minimally considering its on a scale of 6 points). This could be justified by the fact that the skill and knowledge based programs (such as credit analysis and technical analysis) tend to be more difficult and demanding compared to the programs pro-

moting a behavioral changes in the participants (such as time management and negotiation skills). This difficulty arises from the fact that these types of program require more understanding of technical and banking skills, in addition to more input of quantitative skills (e.g., math, statistics, and accounting). Further more, these programs are offered in the evening time since the banks tend not to release technical employees during morning working hours.

It is also obvious that the time effect (Year) has a negative impact on administration and training material. This is due to the following:

The negative time impact on the administration dimension is due to the fact that in the last seven years, due to budgetary constraints, the IBS has not invested in training facilities, which includes training halls, furniture, equipment, and training aids (during this period the total budget of the IBS, which amounts to approximately \$ 3,000,000 p.a. increased only by 2%, most of which was not directed to training facilities).

In addition, the IBS is facing location related problems, since it is located in a narrow space in downtown Kuwait City, with limited parking space and facilities.

On the other hand, the negative impact of time on the material dimen-

sion can be explained by the fact that the IBS has adopted a more specialized training strategy during that period. In which further highly technical, specialized, programs were introduced. These programs tend to be more technical and banking skill oriented and demand the same quantitative skills mentioned above.

Conclusion and Recommendations

This paper has developed and empirically tested a training reaction evaluation model in order to better understand the mechanics of how a trainee forms his/her reaction of a training activity. "Two steps evaluation" exercises are included in the proposed model- daily and final evaluation. The variables of the daily evaluation were relatively highly correlated with the final program attributes, with the exception of the program's administrative arrangement. This means that the trainees overall reaction of a program is not an isolated event that is developed and perceptually formed on the last day of the training program. It is a rational (at least from the trainees' point of view) perception that is related to the day-to-day events of the program (i.e., the daily content and speaker's presentation skills).

The reliability analysis has shown that the variables used from the final

evaluation questionnaire are reliable and consistent in forming each of the four dimensional factors.

The daily content, daily presentation, a time effect variable, and two sets of dummy variables, one related to the nature of program and the other related to the type of trainers, were used in regression models to predict the final evaluation factors (OBJ, REL, METH, MATR, and ADMIN), along with the overall final evaluation.

The regression analysis further substantiated the above. The daily program attributes were found to explain and predict the trainees' reaction. The final program evaluation factors with the exception of the administrative arrangement could also explain the changes in the overall trainees' reaction. This means that the way to achieve higher positive trainees' reaction is through better presentation skills, training material, methods, relevancy, and objectives. Program administrative arrangements is of lesser importance.

The research findings substantiated the importance of daily evaluation in achieving higher trainee satisfaction. Our practical experience in the Institute of Banking Studies also substantiated this importance. The daily evaluation is a mechanism to monitor the program events on a daily basis and therefore overcome any obstacle. Through daily evaluation a training manager can monitor and improve the contents of each session, encourage the trainer to apply an effective

presentation approach, and improve the trainee's perception as it relates to the content of the program, its relevancy, administrative arrangement etc.

However, in order to better utilize the daily evaluation and therefore predict trainee's level of satisfaction of the program (i.e., final evaluation), one should base the daily evaluation questionnaire on more than two variables and/or questions. On the other hand one has to exercise caution in that increasing the number of the questions in the daily evaluation will impact the quality and the trainee's response rate. Further research should investigate the optimal number of session evaluation and the contents of these questions.

Further research should also investigate the effects of gender, trainee's years of experience, education background, subject matter of the program, its duration and timing on the overall program satisfaction. In order to achieve this, the identity of the trainee in the session and final evaluation questionnaires should be revealed. However this could affect the trainee's objectivity in answering the evaluation forms. To overcome this difficulty, we recommend the utilization of controlled experimental research methodology.

Acknowledgements:

The authors wish to thank Dr. Mohammed Al-Mahmeed for aiding in the statistical analysis of this article.

References

- Alliger, G. M. and Janak, E. A. 1989. Kirkpatrick's Levels of Training Criteria; Thirty Years Later. *Personal Psychology*, 42:331-340.
- Al-Khayyat, M. Ridha and Elgamal, A. Mahmoud. 1997. A Macro Model of Training and Development: Validation. *European Industrial Training*, 21, 3:87-101.
- Brinkerhoff, R.O. 1987. *Achieving Results From Training*. San Francisco: Jossey -Bass.
- Camp, R. Richaud, Blanchard, P. Nick and Gregory, E. Huszczo. 1986. *Towards A More Organizationally Effective Training Strategy and Practice*. Prentice-Hall, Englewood Cliffs, NJ.
- Campbell, J.P. Dunnette, M.D., Lawler, E.E. and Wick, K.E. 1970. *Managerial Behavior, Performance and Effectiveness*. New York: McGraw-Hill.
- Clement, R. W. 1982. Testing the Hierarchy Theory of Training Evaluation; An Expanded Role for Trainee Reactions, *Public Personnel Management Journal*, 11:176-184.
- Dionne, Peirre. 1996. The Evaluation of Training Activities: A Complex Issue Involving Different Stakes. *Human Resource Development Quarterly*, 7, (3):279-286.
- Faerman, R. Sue and Ban Carolyn. 1993. Trainee Satisfaction and Training Impact; Issues in Training Evaluation. *Public Productivity And Management Review*, Vol XVI, 3:299-314.
- Foshay, W., Silber, K. and Westgaard, O. 1988. *Instructor's Competencies*. Vol 1; The Standards. Iowa City, Iowa: International Board of Standards for Training, Performance, and Instruction.
- Gagne, R.M. 1977. *The Conditions of Learning*. New York: Holt Rinehart and Winston.
- Goldstein, I.L. 1974. *Training Program Development and Evaluation*. Wadsworth Publishing Co., Monterey, California.
- Goldstein, I. L 1980. Training in Work Organizations. In *Annual Review of Psychology*. Palo Alto, CA: Annual Reviews.
- Goldstein, L. Irwin. 1986. *Training in Organizations; Needs Assessment, Development and Evaluation, Second Edition*, Brooks/Cole Publishing Co., Pacific Grove, California.
- Hamblin, A.C. 1974. *Evaluation and Control of Training*, New York. McGraw-Hill.
- Hicks, W.D. and Klimoski, R.J. 1987. Entry Into Training Programs And Its Effects On Training Outcomes; A Field Experiment. *Academy Of Management Journal*, 30:542-552.
- Holton III, F. Elwood. 1996. The Flawed Four-Level Evaluation Model. *Human Resource Development Quarterly*, 7 (1):5-21.
- Kaufman, R and Keller, J.M. 1994. Levels of Evaluation; Beyond Kirkpatrick. *Human Resource Development Quarterly*, 5:371-380.

- Kirkpatrick DL. 1959a. Techniques For Evaluating Training Programs. *Journal Of American Society For Training And Development*, 13 (11):3-9.
- Kirkpatrick DL. 1959b. Techniques for Evaluating Training Programs; Part 2 - Learning. *Journal Of American Society For Training And Development*, 13 (12):21-26.
- Kirkpatrick DL. 1960a. Techniques for Evaluating Training Programs; Part 3 - Behavior. *Journal Of American Society For Training And Development*, 14 (1):13-18.
- Kirkpatrick DL. 1960b. Techniques for Evaluating Training Programs; Part 4 - Results. *Journal Of American Society For Training And Development*, 14 (2):28-32.
- Kirkpatrick DL. 1976. Evaluation. In Craig RL, (Ed.), *Training and Development Hand Book*:301-319, New York; McGraw Hill.
- Kirkpatrick DL. 1996. Create Ideas Revisited, *Training And Development*, January, 50 (1):54-59.
- Kirkpatrick, LK. November, 1983. Four Steps To Measuring Training Effectiveness. *Personnel Administrator*, 28 (11):19-25.
- Mager, R.F. 1975. *Preparing Instructional Objectives*. Petman Learning Inc., Belmont, California.
- Mahmoud, M. Mohamed. 1991. Descriptive Models Of Student Decision Behavior In Evaluation Of Higher Education. *Assessment And Evaluation In Higher Education*, 16 (2): 133-146.
- Mathieu, J. E., Tannenbaum, S.L. and Salas, E. 1992. Influence On Individual and Situational Characteristics On Measures Of Training Effectiveness. *Academy Of Management Journal*, 35 (4):828-847.
- Munson, S. Lawrence. 1992. *How To Conduct Training Seminars: A Complete Reference Guide For Training Managers and Professionals*, Second Edition., McGraw-Hill, Inc. New York, St.Louis.
- Nadler, Leonar. 1983. *Designing Training Programs: The Critical Event Model*. Addison-Wesley Publishing Co., Massachusetts, California.
- Noe, R. A. and Schmitt, N. 1986. The Influence Of Trainee Attitudes On Training Effectiveness; Test Of A Model, *Personnel Psychology*, 39:497-523.
- Patrick, John. 1992. *Training Research And Practice*. Academic Press, London, San Diego.
- Phillips, J.J. 1995. *Return On Investment - Beyond The Four Levels*. In E.F. Holton III (Ed.), *Academy of HRD 1995 Conference Proceedings*.
- Rothwell, J. William and Kazanaz, H.C. 1992. *Mastering The Instructional Design Process; A Systematic Approach.*, Jossey-Bass Publishers. San Francisco.
- Saari, L.M., Johnson, T.R., McLaughlin, S.D. and Zimmerle, D.M. 1988. A Survey Of Management Training and Education Practices in U.S. Companies. *Personnel Psychology*, 41:731-743.

- Swierczek, W. and Carmichael, L. 1985. The Quantity and Quality of Evaluating Training. *Training And Development Journal*, 39(1): 95 - 99.
- Tannenbaum, I. Scott and Woods, B. Steven. 1992. Determining A Strategy For Evaluating Training; Operating Within Organizational Constraints. *Human Resource Planning*, 15 (2):63-81.
- Tracey, R. William. 1981. *Evaluating Training and Development Systems*. American Management Association, Inc., USA.
- Truelove, Steve (Ed). 1995. *The Handbook of Training and Development*, 2. Fielden House, U.K.
- Wexley, K. N. and Baldwin, T.T. 1986. Post-Training Strategies For Facilitating Positive Transfer; An Empirical Exploration. *Academy Of Management Journal*, 29:503-520.

الملخص

نموذج التقييم التفاعلي ثنائي الإجراء للتدريب والتطوير

عبدالعزیز عبدالمحسن تقي
الهيئة العامة للتعليم التطبيقي

رضا محمد الخياط
معهد الدراسات المصرفية

يعد التقييم التفاعلي للتدريب من أكثر النماذج المستخدمة لتقييم أنشطة التدريب من بين نماذج كير كباترك الأربعة. وحتى يتسنى فهم العوامل المؤثرة في تكوين ردة فعل المتدرب للبرامج التدريبية، فقد تم تطوير واختبار نموذج للتقييم التفاعلي ثنائي الإجراء. تشير نتائج الاختبارات إلى أن التقييم اليومي (الإجراء الأول)، ومتغير الوقت، ومجموعتين من المتغيرات الصورية المتعلقة بتصنيف البرامج، ذات علاقة بالتقييم النهائي (الإجراء الثاني). وهذا يقودنا إلى نتيجة مؤداها أن ردة فعل المتدرب لبرنامج التدريب ليس ناتجا فقط من تفاعله مع أحداث اليوم الأخير للبرنامج، إنما ناتج عن تعرضه للأحداث اليومية للبرنامج ككل. وقد أثبت اختبار تحليل الانحدار ذلك، حيث أثبت الاختبار أن التقييم اليومي للبرنامج متنبأ جيد لردة فعل المتدرب، كما أثبت الاختبار أن متغيرات التقييم النهائي للبرنامج، باستثناء متغير الترتيبات الإدارية، قادرة على تفسير المتغيرات التي تكون الانطباع النهائي للمتدرب عن البرنامج.

Ridha M. Al-Khayyat (Ph.D. University of Southern California, 1991), Head of Training Unit at the Institute of Banking Studies (IBS), also functions as a Core Faculty Member and Researcher at IBS for the areas of organizational behavior, training and development. He is a qualified practitioner of Myers-Briggs Type Indicator. Research Interest: Training development evaluation, psychometrics, computer based training (CBT) and managerial effectiveness.

Abdulaziz A. Taqi (Ph.D. UWIST 1984), Associate Professor of Management, Department of Management, College of Business Studies in Kuwait. Research Interest: Management Training, Organizational Behavior and Innovation.

المجلة العربية للمعلوم الإنسانية

علمية . أكاديمية . فصلية . محكمة

تصدر عن مجلس النشر العلمي - جامعة الكويت

صدر العدد الأول في يناير ١٩٨١

رئيس التحرير: أ. د. عبدالمالك خلف التميمي

الانتراكات

- الكويت: 3 دنانير - ديناران للطلاب - 15 ديناراً للمؤسسات .
- الدول العربية: 4 دنانير للأفراد - 15 ديناراً للمؤسسات .
- الدول الأجنبية: 15 دولاراً للأفراد 60 دولاراً للمؤسسات .

بحوث باللغة العربية والإنجليزية
ندوات - مناقشات - عروض كتب - تقارير

توجه المراسلات إلى رئيس التحرير:

ص.ب: 26585 الصفاة - رمز بريدي 13126 الكويت

هاتف: 4817689 - 4815453 - فاكس: 4812514

e-mail: ajh@kuc01.kuniv.edu.kw

يمكنك الاطلاع على المجلة باللغتين العربية والإنجليزية مع الفهرس على شبكة الانترنت

<http://kuc01.kuniv.edu.kw/~ajh>