

# Is Training Effective? Evaluating Training Effectiveness in Call Centers

Waseem Rehmat  
Iiris Aaltio  
Mujtaba Agha  
Haroon Rafiq Khan

## Abstract

Due to complex, competitive and crucial nature of call center jobs, organizations in services industry are spending more resources than ever on staff training and development. This is the case also in Call Center Representative Training. However, although organizations invest billions of dollars every year in training, no concrete evaluation framework exists to adequately quantify the impact of Call Center Representative (Henceforth CCR) training on actual job performance. Filling this gap, current study attempts to develop a framework to evaluate training programs in the context of call center industry using Kirkpatrick's learning and training evaluation model. Developed framework is then implemented in actual training programs of the case company to develop insights on evaluation of training programs and their limitations. The study is based on actual data of three call centers of a leading Telecom Company in Pakistan. These call centers answer approximately 72 Million calls a year. Study analyzed data of almost 627 CCRs who were trained in 34 different training programs by 18 different certified trainers at three locations. CCR training was selected as research setting because of two reasons. Firstly, high turnover of CCR's in call center industry necessitates frequent and extensive training which makes CCR training a big chunk of resources utilized in call center industry on

training and development; secondly, standardized scrutinizing procedures followed in the call center industry for hiring CCR's enabled and facilitated implementation of training evaluation framework which is suggested in this paper. Data was scientifically recorded for the entire year 2012 and different aspects of training were recorded to ensure that Kirkpatrick model could be applied. By successfully applying Kirkpatrick's learning and training evaluation model, the study developed a framework to gauge effectiveness of training program in call center using Kirkpatrick model. Our investigation of training programs using the developed framework revealed that training programs get very high scores at initial level. Trainees are inclined to rate trainings as excellent at level 1 (Reaction) of Kirkpatrick model but as we go deep with levels (Learning, Behavior) of model, it was identified that effectiveness of training programs deteriorate subsequently. Decline of almost 20% was recorded between the effectiveness of training at Level 1 (Reaction) and Level 3 (Behavior). These results suggest that reaction of trainees is an inadequate measure to evaluate training programs and training programs should be evaluated at a deeper level to get a realistic picture of training effectiveness. Though scope of this study was limited to call center trainings where results at each level of Kirkpatrick model could be gathered objectively, the study opens an interesting and challenging area for management researchers about exploring and improving quality of training programs. It shows the need to study further

this field by developing and implementing effective evaluation models in diverse training fields, specifically in areas such as social and leadership training.

**Key Words:** Training and Development, Learning, Kirkpatrick model, Training Management, workplace effectiveness, Human Resource Development.

## Introduction

Call center industry is growing exponentially. Call centers are vital part of any business because businesses are built around customers and customers want to communicate. They want to tell about their service experiences, issues, complaints and they also want to know about new products, offers and packages that are being launched by businesses. Therefore, organizations while developing their marketing and customer care strategies, consider Call Centers as crucial pillar (Gilson & Khandelwal, 2005).

Though organizations acknowledge the importance of call centers as important pillar of the business, many organizations consider call centers as cost centers because primarily, call centers work as after sales support which does not create any new business unless call center is outbound. On other hand, we do have companies which consider call centers as their profit centers by upselling and cross selling different products when customer is interacting with Call Center Representative (CCR). In either case, businesses want maximum out of their call centers both in term of productivity and quality.

As (Houlihan,2000) points out, most of the work at call centers is managed with the use of technology which determines the pace and volume of work. This system also allows constant monitoring of job and employee performance (Hutchinson, Purcell, & Kinnie, 2000). The call center work environment is characterized as being similar to assembly line production (Taylor & Bayn,

1998) and this creates tough performance criteria's for CCR's. Call center job is considered as one of the toughest job throughout the world resulting in very high turnovers. Estimated average turnover is between 35 and 50 percent (IBIS World, 2008). High stress levels and huge workloads are major contributors in high turnover at call centers. This is the prime reason that almost throughout the year; call centers are hiring resources to fill in resignations.

High turnover and constant hiring usually creates a workforce with unequal skill levels but customers expect same level of services whenever they contact helpline. They need CCRs to be cooperative, friendly, courteous, and attentive with updated knowledge of each and every product, service, and issue. Customers don't care if CCR is new or old and neither should they because it is company's responsibility to ensure that right person is sitting at helpline to facilitate customer. Therefore, with high turnover and constant hiring of new resource, management has to ensure that a standard value has been added in raw resource to meet customer expectations. So, rigorous training is needed to standardize skill levels in the workforce if customer expectations are to be met by call centers.

With heavy investments in executing training programs, the question is no longer "should we train", but rather — is the training worthwhile and effective? So it all boils down to effectiveness of training programs which is done through training evaluation. Problem with currently available training evaluation models is their inability to objectively measure effectiveness of training programs at levels deeper than trainee responses. Though these models present a framework of training evaluation at different levels (i.e. Reaction, learning, behavior), applicability of these models is limited because measuring scales of evaluation are mostly industry specific and highly generalized evaluation models available in literature appear to practitioners as not applicable. Therefore, in the current study, we took the most widely acknowledged training evaluation model (Kirkpatrick model); tailored training programs in our case organization according to evaluation model and implemented the model. This approach to evaluate training programs was inverse to currently suggested and practiced approaches in the field of training evaluation because customarily, evaluation models are implemented on completed training programs whereas in this study, complete training lifecycle was developed in a manner that permitted and supported objective evaluation.

## Training in Call Center Industry

Training is a key strategy for human resource development, generating new skills in people and in achieving organizational objectives. Training can be defined as "the systematic acquisition of skills, concepts, or attitudes that must result in improved performance of the trainee" (Goldstein & Ford, 2002; Aamdor, 2012). Employees/ workforce need to acquire special skills and knowledge appropriate to perform job as per desired standards and training programs are developed to help them achieve those desired targets. Training has many benefits and hence, is becoming a billion dollar industry worldwide. On average, organizations are spending 2 to 2.5 percent of their payroll on training (ASTD, 2005). Same source reveals that U.S. organizations alone spent approximately \$164.2 billion on employee learning/executive education in 2012.

In the call center industry, training programs are developed and executed to inculcate required knowledge, skills and abilities (KSA) in new hires. Training new hires is specifically a very demanding job because there is always high pressure from

call center management to handover staff as soon as possible so that quantitative service levels at call centers may not be compromised. So, the question in this context is, *How to gauge effectiveness of training programs?* Evaluation of the effectiveness of training programs is critical because without it, call centers have no good way to know if CCR will be able to provide standard services to their customers.

Training in Call Center industry is different from training in other organizations because in Call Centers, CCR (Employee) must perform all standard activities with standard accuracy and courtesy at a set level from very first day of the job. Margin of "Trial and error" and "Experiential learning" in Call Center industry is considerably low compared to other industries and business segments where new employees can learn from mistakes and peers.

### Training Evaluation:

Training evaluation is a systematic process of collecting data in an effort to determine the effectiveness and/or efficiency of training programs and to make decisions about training (Brown & Gerhardt, 2002; Brown & Sitzman, 2011). Evaluating training programs is becoming an important issue for training researchers and practitioners (Alliger, Tannenbaum, Bennet, Traver, & Shotland, 1997) because training evaluation is both costly and intensive (Salas & Cannon Bowers, 2001), and evaluation criteria must be psychometrically sound (Alliger, Tannenbaum, Bennet, Traver, & Shotland, 1997).

In the training lifecycle, evaluation phase is usually the most overlooked part. Often, the value of conducting training evaluations is outdone by the necessity simply to gain participants immediate post-training reactions and results of that are sometimes mistakenly viewed as an indicator of whether or not the training was successful overall. In addition, budgetary, and other constraints have caused many trainers and instructional designers to employ standardized, commercially available evaluation instruments. Advantages of using standardized tools are that 1) They are (presumably) validated because they have been used and refined over time and Therefore, the data and feedback they provide is consequently, likewise (presumably) valid. 2) They can be customized, to the extent that many contain questions of open format, allowing the course designer some flexibility of inserting course specific questions and 3) they are relatively inexpensive and readily available thereby allowing the instructional designer to focus mainly on course and curriculum development related concerns. However, there are many disadvantages in using standardized evaluation instruments. Firstly, they present a "one size fits all" approach to training course design in which they assume that each course has relative similarities in its content, style, and expectations. Secondly, they are generally not as comprehensive nor focused on critical content (objective-driven) areas as would be either necessary or desirable and thirdly, they offer little assistance in assessing the longer-term effects of the training.

A valuable alternative to standardized evaluations can be found in designing a customized and systematic approach in which the principal goal is to obtain feedback aimed specifically at a particular program's objectives to determine not only how well the course was initially received but also whether or not it had the desired impact over a sustained period of time.

Existing literature proposes different models for carrying out training evaluation (i.e. Kirkpatrick, 1976; Phillips, 1997; Hamblin, 1974; Tannenbaum & Woods, 1992; Kaufman & Keller, 1994; Holton, 1996). Evaluation approaches used in these models can be loosely categorized into "Goal based" ap-

proaches and “System based approaches”. Various frameworks for evaluation of training programs have been proposed under the influence of these two approaches (Eseryel, 2002). The most influential model for training evaluation with goal oriented approach came from Kirkpatrick whereas under the systems approach, the most influential models include: Context, Input, Process, Product (CIPP) Model (Worthen, Sanders, & Fitzpatrick, 1997); Training Validation System (TVS) Approach (Fitz-Enz, 1994); and Input, Process, Output, Outcome (IPO) Model (Bushnell, 1990). (Eseryel, 2002) provides a comparison between Kirkpatrick model and TVS, IPO and CIPP models which is re-presented in the table below.

Kirkpatrick (1959)	CIPP Model (1987)	IPO Model (1990)	TVS Model (1994)
1. Reaction: to gather data on participants reactions at the end of a training program	1. Context: obtaining information about the situation to decide on educational needs and to establish program objectives	1. Input: evaluation of system performance indicators such as trainee qualifications, availability of materials, appropriateness of training, etc.	1. Situation: collecting pre-training data to ascertain current levels of performance within the organization and defining a desirable level of future performance
2. Learning: to assess whether the learning objectives for the program are met	2. Input: identifying educational strategies most likely to achieve the desired result	2. Process: embraces planning, design, development, and delivery of training programs	2. Intervention: identifying the reason for the existence of the gap between the present and desirable performance to find out if training is the solution to the problem
3. Behavior: to assess whether job performance changes as a result of training	3. Process: assessing the implementation of the educational program	3. Output: Gathering data resulting from the training interventions	3. Impact: evaluating the difference between the pre- and post-training data
4. Results: to assess costs vs. benefits of training programs, i.e., organizational impact in terms of reduced costs, improved quality of work, increased quantity of work, etc.	4. Product: gathering information regarding the results of the educational intervention to interpret its worth and merit	4. Outcomes: longer-term results associated with improvement in the corporation's bottom line- its profitability, competitiveness, etc.	4. Value: measuring differences in quality, productivity, service, or sales, all of which can be expressed in terms of dollars

Table 1. Goal-based and systems-based approaches to evaluation (Eseryel, 2002)

From these and many other models, the most popular and recognized model of training evaluation is Kirkpatrick's (1994) model of training evaluation (Saks & Burke, 2012). Though critiqued for its simplistic approach, this model is preferred for current research because firstly, it provides appropriate goal orientation which is specifically required in call center training and secondly, clear job performance indicators used in call center industry make it possible to implement Kirkpatrick's model with clarity and conviction.

#### Kirkpatrick model:

As per Kirkpatrick model, training can be evaluated at four levels. Level 1 is reactions criteria which evaluates trainees' reactions to a training program. Level 2 is learning criteria, which evaluates the extent to which trainees have learned the training material and acquired knowledge from a training program. Level 3 is behavior criteria and it assesses the extent to which trainees have applied the training at workplace in terms of their behavior and/or performance following a training program. Level 4 is results criteria, which calculates the extent to which

the training program has improved organizational-level outcomes (Kirkpatrick, 1976).

Currently available research on Training evaluation using Kirkpatrick model has mostly been focused on the extent to which organizations evaluate trainings at one or more levels of Kirkpatrick model. For example, few references cited by (Saks & Burke, 2012) suggest that many organizations evaluate reactions and learning, but very few evaluate behavior and results criteria (Blanchard, Thacker, & Way, 2000; Hughes & Campbell, 2009; Kraiger, 2001; Sitzmann, Casper, Brown, & Ely, 2008). Similar results were found in the study made by (Twitchell, III, & Jr., 2000) in which only 31 percent of organizations used behavior measures to evaluate technical training programs, and only 21 percent used results–performance measures. Considering causal relations amongst Kirkpatrick's four levels of training evaluation, interesting study was made by (Alliger & Janak, 1989) in which they found very low correlations between reactions and the other three criteria (Learning, Behavior and Results) and slightly larger (but still relatively low) correlations among learning, behavior and results criteria.

Therefore, though previous studies have used Kirkpatrick model to assess training efficacy, not enough research evidence was found where complete lifecycle of training was used to gauge effectiveness of training programs in general and in the call center industry specifically. Filling this gap, we developed a framework of complete training lifecycle using Kirkpatrick model and implemented the model on results of training evaluation. In this context, the research aimed at answering following two research questions.

a) How can we develop a framework to evaluate training programs in the context of call center? What can we learn using the Kirkpatrick's model in evaluation?

b) How effective are call center new hire training program at different levels of Kirkpatrick's model?

Aim of the study therefore, was to contribute in current discourse of training evaluation via appraising the effectiveness of the evaluation process and by providing a framework of complete training lifecycle in which evaluation is embedded in the planning and execution phase. Though scope of this study was limited to call center trainings where results at each level of evaluation could be gathered objectively, the study opens an interesting and challenging area for management researchers about exploring and improving quality of training programs by developing and customizing evaluation models for diverse training fields such as soft skills and leadership training.

#### Methods and Datasets:

This research was carried out in a Telecommunication company in Pakistan. 34 New hire training programs were analyzed for this research. All trainings that have been conducted in the organization in year 2012 have been analyzed. Data set that has been used for research comprised of 627 CCRs who actually went through new hire training. Their performance was recorded for respective next couple of months to capture the results as per Kirkpatrick model. New hire training has been chosen for this research mainly because new hires do not carry any specific cultural or political biasness/ inclinations and it was assumed that choosing new hires would provide true results while evaluating post training behavior. 34 complete actual new hire training programs were monitored and recorded to analyze effectiveness of training program at each level of Kirkpatrick model. It took almost 12 months to generate this data. Basic processes and phases of new hire training program



are explained below to get understanding of the entire structure of training.

**Hiring Process:** Basic requirement for entry level hiring (new hires investigated in this research) at the organization under study is graduate level studies and age limit of  $\leq 28$  years. Company does not advertise jobs of CCR's as it is a continuous process and usually existing employees refer other people. There is a CV bank in which all the resumes are dropped and are pulled off as and when needed. At any given time, there are thousands of CV's in CV bank. Call center coordinator maintains the record of all resumes. Hiring process comprises of following steps.

- CV Short listing- For all the CVs initial screening is done; basic qualification, prior experience and other key parameters are keenly observed to short list CVs.

- Prioritizing CVs- As basic criteria is very simple and there are hundreds of short listed CVs, so next step is prioritize CVs. Persons having prior experience in similar jobs are given priority.

- Telephonic interview-Once CVs are prioritized, panel of team leaders start calling prospects, to Confirm their availability, Gauge basic telephone skills and to check voice quality on telephone. They record their feedback in prescribed format against each interviewee. All candidates succeeding the telephonic interview phase are called in for assessment center.

**Assessment Centers:** There was always a conflict between training team and management regarding quality of intake at call centers. Trainers usually complain that hired CCR's do not possess basic capabilities that are required for job and hence, cannot be trained adequately for the job. To resolve this issue, a committee was developed comprising of all the stake holders including call center management, quality assurance team and training team. This committee came up with solution of "*pre-hiring assessment centers*". After thorough deliberation, following competence areas were established and agreed to be gauged during assessment centers.

- a) IQ test
- b) System handling skills
- c) Communication skills
- d) Customer Focus

Each of these competence areas have been established after carefully looking at job description of CCR's. As part of their job they will be interacting with customers and therefore, they need to be extremely customer focused. Communication is another key competence required for Call Center job. Similarly, every task that they perform during job is on system and therefore, they must be proficient in using systems. Certain IQ level is also needed to comprehend customer issues, understand logics, analyze problems and give solutions to customers.

All these competence areas were agreed by all the stake holders' i.e. Call center management, QAST (Quality Assurance and Standardization Team) and VP Customer Care. Hiring through assessment center also mitigated the limitation of Kirkpatrick model where literature claims that right selection of candidates for training programs will affect the results of training. So, by selecting human resource through assessment center, it was assured that right people were selected for the job.

Assessment centers basically ruled out the possibility of wrong candidate selection. This ensured that candidates with required skill set were chosen and sent to new hire training program. This resulted in nullifying the limitation of Kirkpatrick model where some literature claims that result of training depends on the skillset of trainees. Detail of each competence area and how

it was gauged in assessment centers is mentioned below.

**IQ Test:** IQ test comprised of two parts. First part contained 10 sequence related questions, and second part was comprehension to gauge if participant was able to use written material regarding product knowledge that will be used in trainings and afterwards on job. Sample questionnaire used to gauge IQ of participants is shown in Appendix 1.

**System Handling Skills:** System handling skills were gauged by asking participants to perform different activities on computer. It was again a simple sequence of tasks that was performed on computer. Evaluators had assessment sheets and while participants used systems, evaluators assessed each person's competence regarding system handling. Each participant was assessed by at least two evaluators to ensure that the marks given to participant were correct. Each participant was scored on a scale of 1-10. Instructions sheet used for system handling skills is attached in Appendix 2.

**Communication Skills:** Communication skills were gauged in two dimensions. One was speaking skills and other was listening skills. For speaking skills, participants were asked to speak on any topic for two minutes and as they spoke, expert evaluators evaluated them on different parameters of speaking skills such as opening of presentation, body language, selection of words and closing of presentation. Score by each evaluator were then summed up to get the total score for participant.

Second area of communication skills gauged was listening skills. Two recorded calls were played while participants listened with a blank paper and pencil in their hand. Calls were actual conversations between customer and CCR. In these calls, usually customer is asking for something or is discussing his/her issue and CCR is providing relevant information to customer or is trying to resolve the customer issue and providing appropriate remedies.

Recorded call was played, participants carefully listened to the call and they could take notes if they wanted to and once call was ended. Evaluators served question paper regarding that specific call. Each call had five questions and each question was worth one point. Listening skills had 10 marks in total and speaking skills was also of 10 marks. So, participants got scores out of 20 in communication skills category.

**Customer Focus:** Gauging if candidate was customer focused or not proved to be a tricky job. A questionnaire was developed to get the feel about aptitude of participant towards serving customer. This questionnaire basically comprised of daily routine job scenarios where one could observe if candidate was customer oriented or not. There were 10 questions and each question carried one mark. So participant's score was calculated out of 10.

**Overall Score:** Once all the competence areas were gauged, scores were calculated and compiled on an Excel sheet and Candidates were ranked according to the score they achieved in the assessment. Candidates failing to meet passing criteria were given feedback on weak areas and their application for job was declined with remarks.

Table 2 (p. 8) gives the summary of weightages against each competence. There is another column in same table which is "*Passing criteria*". As all competences are necessary to perform the job of CCR so s/he had to get a passing score in each competence which is mentioned in front of each competence in the table.

Apart from getting passing scores in each above criteria total score to pass the overall assessment needed to be greater than 74%. Compiled results were sent to call center management to continue further with hiring process. Furthermore, Candidates

Competence Area	Weight	Passing Criteria
IQ Test	20%	70%
System Handling Skills	40%	80%
Communication skills	20%	80%
Customer focus	20%	80%

Table 2: Assessment center competencies

Scores	Category
95% - 100%	A+
85% - 95%	A
75% - 85%	B
<75%	U

Table 3: Assessment Center Overall Scores Ranking

were given rating depending on their performance in assessment center. Table 3 summarizes the criteria to rank candidates based on their overall score.

Once candidates got cleared from assessment centers, they were called in for interview with line manager. Line manager interviewed applicants for approximately 10 minutes on defined format and recorded her feedback. After approval and clearance from line manager, applicants were called for final Interview. This interview was a panel interview in which Head of department along with HR representatives interviewed applicants. This was the final stage of hiring and candidates passed through this stage were selected for job and given offer letter.

On acceptance of offer letter, CCRs were requested to join in training. For our study, this extensive and standardized hiring criteria was important and useful because through the filtration mechanism, we could safely assume that selected candidates possessed standard pre-training skills and competencies and we could predict that post training evaluations and results would be training dependent and not based on inequality in trainee selection.

#### CCR Performance Standards:

CCR performance is closely monitored in call centers. There are certain KPIs against which CCRs are being monitored, incentivized, appraised and promoted. As our study evaluated post training behavior (Level 3 of Kirkpatrick model) against these KPIs, it was very important to understand call center performance indicators. Therefore, each performance indicator used for performance evaluation is mentioned below in detail.

**Quality scores:** 20 customer interactions (calls) of each CCR are evaluated every month. Quality Assurance team evaluates 10 calls against desired parameters and each call is given rating out of 100%. Afterwards, Quality Assurance team calls out 10 customers who interacted with particular CCR and asks his/her feedback about the call against defined parameters. Cumulative score of these 20 calls is the overall quality score of that particular CCR. Major part of CCR performance is to ensure that he/she gets maximum score in this parameter.

**Productivity Score:** Productivity is second most important parameter of CCR performance. Productivity basically is the measure to see how CCR performed in his/her 8 hours shift i.e. Did CCR logged in on time? Did CCR catered customer calls within defined time? Took his/her breaks properly? CCRs are supposed to handle each call in 120 seconds on average. CCRs are given three breaks in 8 hour shift as per pre-defined schedule. All these parameters are captured automatically in system and system generates Productivity score of each CCR for entire month.

**Quiz Score:** There are lot of things to remember about different products that are being offered by company because each day, company launches different offers and packages to attract and facilitate customers. As contact point for customers, CCRs are supposed to know about all the products that are being offered by company. To ensure this, there is a monthly quiz regarding different offers/packages/scenarios/promotions etc. that is conducted for each CCR. Quiz is also conducted by an automated system and questions are randomly chosen from system from a large database of questions. Each quiz contains 10 questions and system generates report of each CCR performance in quiz every month.

**Attendance:** An uninformed leave highly affects service levels of that day; CCRs attendance is strictly monitored and is part of CCR performance evaluation. Informed and approved leaves do not affect CCR performance but uninformed absence from job would severely affect CCR performance in that month. Weightage of each performance indicator for CCR at job is mentioned below table 4.

**ABU Report:** All these KPIs are summed up as per their assigned weightage in one report called ABU. ABU basically grades performance as "A" in case performance is good, "B" in case of satisfactory performance and "U" grade for unsatisfactory performance. Detail rating is illustrated in Table 5.

KPI	Weightage
Quality	55%
Productivity	30%
Quiz	5%
Attendance	10%

Table 4: Call Center Representative KPIs

ABU Grade	Percentage Score
A+	>=95%
A	>=90% < 95%
B	>=85% < 90%
B-	>=80% < 85%
U	<80%

Table 5: ABU Scores

**Training Design:** Once candidates are hired, they are requested to join in the training program. It is 14 days training program developed to equip new hires with necessary KSA (Knowledge, Skills and Attitudes) that are required to perform the job. New hire training process was developed by the training experts and in coordination with all stake holders. In the observed organization, each training program is designed to serve at least one business objective and Keeping that business objective in mind entire training program is developed. As potential trainees observed in our study were customer service department employees who were supposed to handle customer queries and concerns over the phone, following business objectives were paramount while developing training modules.

- At the end of training, trainees must have adequate product knowledge to handle customers effectively. They must be able to provide information that is required by customer.

- Trainees must exhibit customer handling skills. They must be polite and courteous towards customers in their job.

- By the end of training, trainees must be able to use relevant systems to handle customer queries and concerns.

To deliver on desired objectives, new hires training program was divided into two prominent areas. Initially, there was 10 days training in class room. These training rooms were equipped with necessary systems (with necessary systems). All the prod-

uct knowledge, system knowledge etc. was taught to trainees in those 10 days. In last 4 days, trainers placed these new hires on job and observed how they performed while mentoring and assisting them at work. In 14 days program certain number of quizzes were incorporated. Different activities like mock decision making situations, role plays, and team competitions were added in modules to ensure that participants learn as much as possible.

After module development and finalizing the complete training module hour by hour, trainee manual was developed. That was a simple handbook for trainees which contained all the relevant knowledge that was to be taught during the training program.

**Training Execution:** Once module and content was developed for trainings, the next step was to execute trainings. While executing training, certain criteria's were considered like arranging appropriate room, having appropriate seating arrangement etc. During training, participants were given several assignments, quizzes, presentation, notes etc. to keep them involved. Two trainers were always present with trainees and usually they divided different topics among themselves to train participants turn by turn. Entire batch was the responsibility of those trainers. Training team used mix and match of different trainers in different trainings to ensure that participants could get the mix of different trainers and at the same time different trainers could work together to improve their training skills further. Several parameters were gauged in training program to ensure that new hires might perform on job as per desired standard set already during the planning phase of training. Training performance of trainees was measured in the following manner.

- Attendance (Weightage 10%) - Popper record of attendance for each day was kept to ensure that new hires don't miss any topic.
- Training Quizzes (Weightage 10%) - Quizzes were conducted during training to gauge the learning of trainees.
- Training Participation (Weightage 10%) - Trainees were encouraged to actively participate in training; their participation was also gauged and recorded.
- Training Projects/Assignments (Weightage 10%) - Several projects/assignments were given to trainees in training program. These assignments and projects were evaluated and recorded to calculate final score of trainees.
- Final Quiz (50%) - At the end of training, grand quiz was conducted.
- On job observations - (Weightage 10%) - Measured in last 4 days of training.

At the end of training, all above parameters were summed up as per their defined weightages to get the final score of individuals which was later used to measure training effectiveness at level 2 (Learning). At the end of training, scores were compiled and shared with management. Trainings were concluded with a 10 question feedback form filled by each participant. Feedback form was later used to gauge Level 1(Reaction) of training.

#### Training Evaluation using Kirkpatrick model:

Kirkpatrick model presented four levels to actually gauge the effectiveness of training. As this research focuses to implement Kirkpatrick model levels for new hire training program to establish effectiveness of training program, each level of Kirkpatrick model was mapped to different activities at each stage of trainings in the following way.

**Level 1- Reaction-** Reactions is basically the immediate feedback of participants after the training. To calculate

Training feedback (Reaction) of each new hire training batch, feedback forms were used. Feedback form had 10 questions and participants ranked each questions against three given options (Excellent, Good and Needs Improvement). To calculate the result from feedback form, Each excellent that was selected by participant was considered as 5 marks, for good, 3 marks were assumed and Needs improvement resulted in 0 marks for particular question. All the scores against excellent, goods and need improvement were calculated and divided by 50 to get the quantified result of feedback. Average score of Feedback by all participants was considered as the Training feedback (Reaction) to training program.

Feedback form contained couple of questions regarding the training facility and required system support whereas most of the questions were related to capabilities of trainer i.e. was s/he knowledgeable, were the sessions interactive etc. Last portion of feedback asked for suggestions/ideas to improve training program and participants were encouraged to write any improvement they want to suggest.

On 14th day of training, training manager visited the new hire training batch, Trainers left the training room to ensure that feedback was transparent, new hires were given feedback forms and they recorded their feedback on given forms.

Once feedback forms were filled, training manager collected them and gave them to quality assurance team. Quality assurance team entered feedbacks in excel sheet and training results were calculated by compiling all the feedbacks. These results were later shared with call center and training team's management.

**Level 2 – Learning-** Level 2 of model is "Learning" which basically tries to capture how much actually trainees have learnt during the training. One can define it as techniques, knowledge and abilities actually acquired by trainee due to training.

To get this Level 2 evaluation of new hire training, different parameters that were used during training were analyzed. Final result comprised of all the learning that has happened at trainee level because this score was derived from Grand quiz which covered everything taught during the training program. So to gauge effectiveness of training at Level 2, Final training results of participants were used.

**Level 3 – Behaviors-** Level 3 of Kirkpatrick Model is about monitoring actual behaviors that new hired CCRs demonstrated on the job after training. Level 3 (Behaviors) evaluation was generated using ABU, the performance report generated by Quality Assurance Team. As already described, each CCR on floor is evaluated against several parameters and combination of all these scores is the performance score of CCR. After one week of joining floor (Job), quality assurance team starts evaluating new hires performance.

The Quality scores of ABU were considered for gauging training effectiveness at level 3 because other parameters like attendance and productivity were not directly related to effectiveness of training. Additionally, training was never designed to make CCRs more efficient in productivity (Attendance, break management etc.). Quality score basically comprised of new hire actual behavior on calls with customer. Therefore, Quality Score of new hired CCRs in first month ABU after joining the floor was considered as level 3 evaluation of this study.

**Level 4 – Results-** Level 4 of Kirkpatrick model talks about the actual results that affect the organization as a whole on a bigger scale. This is the most complex level of Kirkpatrick model and it needs certain sets of data which organizations don't always have. For new hire training, this can be calculated from number of calls answered by each new hire in a month. Rev-



enue/cost of calls where quality of CCR was good can be considered as return on investment whereas the cost of calls where CCR failed to meet quality standards can be considered as lost opportunity. However, due to complexity of this level, lack of data and to rationalize the scope of this research, level four of Kirkpatrick model was not implemented in this study. Table 6 below summarizes parameters used to capture data from training sessions to implement Kirkpatrick model.

Level 1- Reaction	Feedback of trainees at the last day of training.
Level 2- Learning	Final results of trainees at the end of trainings (Aggregate of quizzes, participation, attendance and observations)
Level 3- Behavior	Departmental evaluations after one month of job
Level 4- Results	Not in the scope of this research

Table 6: New hire training results mapped as per Kirkpatrick Model

**Results and Conclusions:**

34 new hire training programs were analyzed. These were all the new hire trainings conducted during 2012 in case organization. There were 627 participants who were trained in these 34 programs and results of training against each level were measured as previously defined. Table 7 summarizes the results of effectiveness of training at each Kirkpatrick level.

Kirkpatrick Level	Research parameter	Scores
Reaction	Training Feedback	96%
Learning	Training Score	86%
Behavior	Job evaluation score	77%

Table 7: New hire training Results at different Kirkpatrick Level

So if we look at reaction of training which was previously being used as an indicator of training success by organization, we can say that training program was 96% effective. At “learning level 2 of Kirkpatrick model, participants score had declined to 86%. It was also observed that for each new hire training batch, “Learning” was always less than the score that training program had achieved at “reaction” level. Third level in Kirkpatrick model was about the actual behaviors depicted by participants of training program on job. The nationwide score of evaluation at this level revealed the fact that training was actually only 77% successful . Figure 1 summarizes the finding of effectiveness at each level.

After applying Kirkpatrick model to new hire training program, it can clearly be inferred that participants reaction to training is highly inflated and if organization or

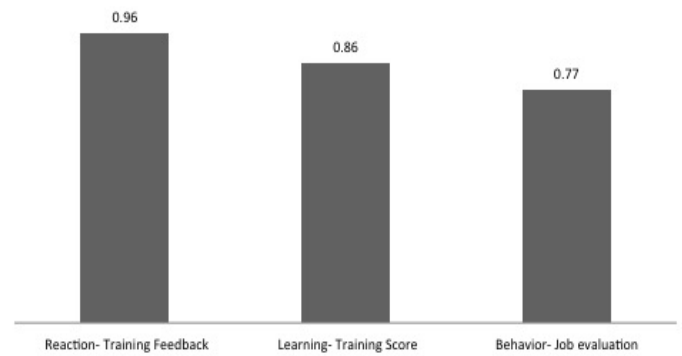


Figure 1: Training effectiveness at each level of Kirkpatrick model

training team is relying only on this single parameter to calculate effectiveness of training than this can mislead to effectiveness perception which in actual is not there. It was observed that for every training batch, training feedback is always beyond 90% which depicts that during training, trainees are impressed by the personality of trainer or they are shy to give negative remarks or low scores due to different reasons. At the end of program they rate training as highly successful.

Figure 2 shows the training feedback of participants against each training program. It is evident that training team had achieved a very healthy feedback on training. Participant’s initial reaction to training was very good.

At next step when we looked at learning’s of training program by calculating the actual performance of trainees in the training, we noticed same trend. Mostly, the average score of each batch was more or less same. There are new hires who get very high score in training and there are few who don’t get very high scores but minimum score needed to pass training was 80% so the range was between 80%-100% as far as individual performance is considered, this was averaged as 86% from all 34 sessions. This level of evaluation suggested that training program was 86% effective with a decline of 10% against 96% effectiveness at reaction level.

Next step of Kirkpatrick model showed some interesting results. Figure 4 (p. 11) is about the results captured at behavior level for all batches. So, the training which was considered as 96% effective was actually only 77% effective at Behavior level which was almost 19% less than actual scores recorded at reaction level. So, answering RQ-2 (How effective are call center new hire training programs at different levels of Kirkpatrick’s model?), analysis revealed that organizations that are considering reaction as the only parameter to gauge effectiveness of training can be highly misled by results as reactions of trainees after end of training do not reflect the true and complete picture of actual training effectiveness.



Figure 2: Training feedback - Reactions



Figure 3: Training score - Learning's



Figure 4: 1st job evaluation score - Behaviour

## Conclusions & Recommendations:

By using the framework and mapping of call center evaluations in Kirkpatrick model in this study, organizations can measure effectiveness of their training programs. Furthermore, analysis in the study revealed that organizations that are considering reactions as the parameter to gauge effectiveness of training can be highly misled by results as Reactions of trainees after end of training are usually highly inflated about trainings. Gauging effectiveness at Level 2 of Kirkpatrick model gives comparatively reasonable results but actual results are what trainee participant delivers on job. Therefore, organizations must strive to capture training results at this level which is the 3rd level of Kirkpatrick model.

At abstract level, we endeavored to contribute in the current discourse of training evaluation by appraising the effectiveness of the evaluation process and by providing a framework of complete training lifecycle in which evaluation is embedded in

the planning and execution phase. Though scope of this study was limited to call center trainings where results at each level of evaluation could be gathered objectively, the study opens an interesting and challenging area for management researchers to exploring and improve quality of training programs by developing and customizing similar evaluation models for diverse training fields such as soft skills and leadership training.

## Limitations & Future Research

This research has been conducted on trainings where results at each level of Kirkpatrick model could have been gathered but the trainings where objective is to inculcate only soft skills such as Communication skills, leadership Training etc., implementing Kirkpatrick model can be difficult. It would be an interesting study to find data points or ways to quantify purely subjective themed trainings.

## References

- Aamodt, M. (2012). *Industrial/organizational psychology: An applied approach*. Cengage Learning.
- Alliger, G., & Janak, E. (1989). Kirkpatrick's levels of training criteria: thirty years later. *Personnel Psychology*, 331-342.
- Alliger, M., Tannenbaum, I., Bennet, W., Traver, H., & Shotland, A. (1997). A meta analysis of the relations among Training criterias. *Personal Psychology*, 341-358.
- ASTD. (2005). *State of the industry Report 2005 Executive Summary*. Alexandria, VA.: American Society for Training and Development.
- Blanchard, P. N., Thacker, J. W., & Way, S. A. (2000). *Training evaluation: perspectives and evidence from Canada*. International Journal of Training and Development, 295-304.
- Brown, & Gerhardt. (2002). Formative evaluation: an integrative practice model and case study. *Personnel Psychology*, 951-983.
- Brown, K. G., & Sitzman, T. (2011). *Training and Employee Development for Improved Performance*. In *Handbook of Industrial and Organizational Psychology* (pp. 469-503). American Psychological Association.
- Bushnell, David S. "Input, process, output: A model for evaluating training." *Training and Development journal* 44.3 (1990): 41-43.
- Camp, R., Blanchard, P., & Huszycz, E. (1986). *Toward a more Organizationally effective Training Strategy and practice*. Engle



- wood Cliffs, NJ: Prentice-Hall.
- Fitz-Enz, J. (July, 1994). Yes...you can weigh training's value. *Training*, 31(7), 54-58.
- Gilson, K., & Khandelwal, D. (2005). Getting more from Call Centers. *The McKinsey Quarterly*, November 22, 1-8.
- Goldstein, & Ford. (2002). *Training in Organizations*. Belmont, CA.: Thomsan Wadsworth.
- Hamblin, A. (1974). *Evaluation and Control of Training*. Maidenhead: McGraw-Hill.
- Holton, E. F. (1996). The flawed four-level evaluation model. *Human resource development quarterly*, 7(1), 5-21.
- Houlihan, M. (2000). Eyes wide shut? Querying the depth of call center learning. *Journal of European Industrial training*, Vol 24, 228-240.
- Hughes P D, & Campbell, A. (2009). *Learning and Development Outlook 2009: Learning in Tough times*. Ottawa: The Conference Board of Canada.
- Hutchinson, S., Purcell, J., & Kinnie, N. (2000). The challenge of the call center. *Human Resource Management International Digest*, Vol 8, 14-18.
- IBISWorld. (2008). *Telemarketing and Call Centers in the US*, 56142. Santa Monica, CA: IBIS World.
- Kirkpatrick, D. (1976). *Evaluating of training*. New York: McGraw-Hill.
- Kraiger, K. (2001). *Creating, Implementing, and Managing Effective Training and Development: State-of-the-Art Lessons for Practice*. Jossey-Bass.
- Kaufman, R., & Keller, J. M. (1994). Levels of evaluation: beyond Kirkpatrick. *Human Resource Development Quarterly*, 5(4), 371-380.
- Ostroff, C. (1991). Training effectiveness measures and scoring schemes: a comparison. *Personel psychology* Vol 44 No. 2, 353-74.
- Phillips, J. J. (1997). *Handbook of training evaluation and measurement methods*. Routledge.
- Saks, A., & Burke, L. (2012). An investigation into the relationship between training evaluation and the transfer of training. *International Journal of Training and Development*, 16(2), 118-127.
- Salas, Eduardo, and Janis A. Cannon-Bowers. "The science of training: A decade of progress." *Annual review of psychology* 52.1 (2001): 471-499.
- Sitzmann, T., Wendy J. Casper, Brown, K., & Ely, K. (2008). A Review and Meta-Analysis of the Nomological Network of Trainee Reactions. *Journal of Applied Psychology*, 280-295.
- taylor, P., & Bayn, P. (1998). An Assembly line in the head: work and employee relations in the call center". *Industrial Relations Journal*, Vol 30 No. 2, 101-117.
- Twitchell, S., III, E. F., & Jr., J. W. (2000). *Technical Training Evaluation Practices in the United States*. *Performance Improvement Quarterly*, 84-109.
- Worthen, Blaine R., James R. Sanders, and Jody L. Fitzpatrick. "Program evaluation." *Alternative approaches and practical guidelines* 2 (1997).

## Authors

**Waseem Rehmat**, (M.Sc. Econ) is a doctoral candidate at Jyväskylä University School of Business and Economics (Finland). Waseem Rehmat possesses over 12 years of professional experience in Training and Development. His research interests include futuristic studies in adult learning, Constructivist learning at workplace, Post training community development, Peer mentoring, Dynamics of online learning and constructivist executive training.  
E-mail: waseem.rehmat@jyu.fi

**Iiris Aaltio**, PhD, is Professor at the Jyväskylä University School of Business and Economics in Finland. Her areas of study are organizational culture, leadership, as well as gender and diversity in organizations. Her recent publications include articles in journals like: *Equality, diversity and inclusion*, *International Journal of Intercultural Relations*, and *Leadership & Organization Development Journal*. She is guest editor in the becoming special issue about "Ageing, Socio-economic Change, Challenges and Potentialities", in the *International Journal of Work Innovation*.  
E-mail: iiris.aaltio@jyu.fi

**Mujtaba Hassan Agha**, PhD, is an Associate Professor at Mohammad Ali Jinnah University (Islamabad Campus) Pakistan. He did his PhD in Industrial Systems Engineering from University of Toulouse (France) and holds an MBA from National University of Science and Technology (Pakistan). He has more than ten years of professional and teaching experience. His research interest include use of Quantitative and Statistical tools to address Operations and Supply Chain Management Problems.

**Haroon Rafiq Khan**, MBA, possesses 10 years of professional experience in customer service trainings, business intelligence and customer experience management in the Telecom industry. His analytical approach towards business processes, trainings, systems and strategy has been rewarded with various recognitions throughout his professional career. His current role involves driving the customers experience and business transformation at one of the top telecom operator's in Pakistan.

## APPENDIX 1

CALL CENTER ASSESSMENT CENTER: IQ & Reading Comprehension Test

NAME	
DATE	
TOTAL QUESTIONS	
ALLOCATED TIME	
Note: Encircle the correct answer and there is no negative marking .Overwriting not allowed.	

### PART 1: IQ Questions

- 1- Sequence 4, 7, 9, 2, 6, 8 read backwards as 8, 6, 2, 9, 4, 7?  
 a- True  
 b- False
- 2- Choose the word most similar to "Trustworthy":  
 a- Resolute  
 b- Tenacity  
 c- Relevant  
 d- Insolent  
 e- Reliable
- 3- How many alphabets are in between Alphabet K and P?  
 a- 6  
 b- 3  
 c- 7  
 d- 4
- 4- If Aslam looks at the mirror and touches his right ear, the mirror will show him touching his left ear.  
 a- True  
 b- False
- 5- The word revolver reads exactly the same back to front  
 a- False  
 b- True
- 6- Which one of the five is least like the other four?  
 a- dog  
 b- Mouse  
 c- Lion  
 d- Snake  
 e- Elephant
- 7- Which number should come next in the series? 1 - 1 - 2 - 3 - 5 - 8 - 13  
 a- 8  
 b- 13  
 c- 21  
 d- 26  
 e- 31
- 8- If you rearrange the letters "ANKIO" you would have the name of \_\_\_\_\_?  
 a- City  
 b- Animal  
 c- Mobile  
 d- River
- 9- Find the next letter in the following series a, c, f, j, o, \_\_\_\_  
 a- Z  
 b- V  
 c- U  
 d- X
- 10- Find the missing number  
 1 4 3  
 5 9 4  
 4 5 \_\_\_\_  
 a- 1  
 c- 10  
 d- 20  
 e- 9

### PART 2: READING COMPREHENSION TEST

Note: Read the below paragraph, please answer all questions.

Jamal works for a company in Lahore, Pakistan. He is a customer service representative. He gets up at six o'clock each workday. He drives to work and begins his job at eight o'clock. He speaks to people on the telephone to help them with their banking problems. People telephone the bank to ask questions about their accounts. He doesn't give information about accounts until people answer a few questions. Jamal asks callers their birth date, the last four digits of their credit card number and their address. If a person gives incorrect information, Jamal asks him to call back with the correct information. He is polite and friendly with everyone. He has lunch in a park next to his office. He returns home at five o'clock in the evening. After work, he goes to the gym to work out. He has dinner at seven o'clock. Jamal likes watching TV after dinner. He goes to bed at eleven o'clock at night.

Instructions: Choose/Circle the correct answer to these questions based on Jamal's Day.

- Q1: What does he do during the day?  
 • He helps customers on cash counter in a bank.  
 • He helps customers on the telephone.  
 • He helps customers outside.
- Q2: What does he do to check information?  
 • He asks people some questions.  
 • He tells them to call later.  
 • He asks to see documents.
- Q3: What does he do if the information is incorrect?  
 • He gives banking account information.  
 • He asks the callers to call back with correct information.  
 • He transfers the call to his seniors.
- Q4: How is the behavior of Jamal on the job?  
 • He is unfriendly and helpful.  
 • He is funny and helpful.  
 • He is polite and friendly.
- Q5: Where does he eat at lunch?  
 • At work.  
 • In a park near work.  
 • At home.

## APPENDIX 2

Instructions for System Quiz

Time Allowed: 10 Minutes

- Please open Google page.
- Search ABC company Website.
- Search for vision of ABC
- Search Prepaid Packages
- Describe Prepaid Package 1
- Copy the vision into a Word Document.
- Now type the vision by yourself. (Copy not allowed here, you need to type in blue color).
- Save the document on your desktop giving it your name.
- Send this document using your email account to imran.nabi@abc.net
- Let the observer know once you are done.