Prof. Jonathan Passmore's Publications Library:

This paper is the source text for the published paper or chapter. It is made available free of charge for use in research. For the published version of this paper please visit the publisher's website. Access to the published version may require a subscription or purchase.

Author(s): Passmore, Jonathan

Title: SOAP–M: A training evaluation model for H.R

Year of publication: 2012

Journal / Source: *Industrial and Commercial Training*

Citation: Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

SOAP- M: A training evaluation model for HR

Abstract

Purpose: The purpose of this paper is to critically review existing models for evaluating organizational HR interventions and develop a practical but robust model for use by practitioners and researchers.

Design: A literature review was undertaken of existing models developed to evaluate organizational HR interventions. Based on these a new model was developed.

Findings: The paper suggests that many of the existing models are either outdated or lack practical focus. The SOAP-M model offers five levels for evaluation, four suitable for HR professionals and a fifth level for researchers.

Research limitations: The paper is conceptual and the model needs to be tested by organizations to assess whether this has utility for HR practitioners.

Practical implications: The model offers a framework for conducting evaluations of organizational interventions and encourages HR practitioners to improve their practice by conducting robust evaluations of the interventions they use.

Originality: The paper offers a new and previously unpublished model which could help HR practitioners improve their use of evaluation.

Key words: *Training, training evaluation, evaluation model, Phillips ROI, Kirkpatrick, CIRO, CIPP Model, IPO Model*

Introduction

Current economic globalization, accelerated by the removal of tariff barriers, the reduction in transport costs, the boom in information and communication technologies, and the internationalization of investments, is drastically changing the scenario in which the world's socio economic players perform (Pellicer, Yepes & Rojas, 2010, Bold, 2011). Nowadays, organizations are challenged to remain competitive and to adapt to the new regulations that are being imposed by the EU and global bodies.

The changing marketplace, with empowered workers and technological advancements, has created an environment where change has become a constant factor. In response, organizations have responded by seeking better returns from their investment in employees. Specifically there has been an increase in the past in the use of competencies frameworks managing to help assessment and seeking higher levels of skills and performance.

The Human Resources functions within global organizations have responded to these demands to deliver improved people outcomes. People management has been recognized as a critical component in knowledge based economies (Kumpikaitė, 2007). This has seen a stronger interest in evaluating each HR interventions to assess what added value it delivers to the business (Kaufman, Keller & Watkins, 1995).

Over the past years, a number of human resource training evaluation models have been developed as tools to help find the dimensions or factors to be considered in evaluation effectiveness (Tzeng, Chiang & Li, 2007). This has resulted in a range of models available for organizations to consider when undertaking HR evaluations such as training or coaching.

A brief critique of training evaluation models

The most popular and widely known approach to the evaluation of training is Kirkpatrick's framework. The model has served as the primary organizing design for training evaluations in organizations for over thirty years. Kirkpatrick identifies four categories of measures: reaction, learning, behavior, and results (Kirkpatrick, 1979).

Level one includes assessment of training participants' reaction to the training program, especially assessment of affective responses to the quality or the relevance of training. This has been incorporated by most organizations into the frequently used training evaluation questionnaire or 'happy sheet'. Level two, learning measures, is defined as quantifiable indicators of the learning that has taken place during the course of the training. Level three, behavior outcomes, addresses either the extent to which knowledge and skills gained in training are applied on the job or result in exceptional job-related performance. Finally level four, outcomes are intended to provide some measure of the impact that training has had on broader organizational goals and objectives (Alliger & Janak, 1989; Bates, 2004).

Critics have highlighted a series of criticisms of the Kirpatrick's model (Bates, 2004). Guerci and colleagues have suggesting that the four levels of evaluation that it proposes lead to an excessively simplified vision regarding the effectiveness of training, particularly because it does not consider the influences of the organizational context (Guerci, Bartezzaghi & Solari, 2010). A second criticism is based on the causal relations between the levels of evaluation. According to the model it is not possible to achieve positive results at top levels if this does not occur at lower levels (Alliger & Janak, 1989). There is limited published evidence to support this. A third criticism of the hierarchical model is the unitary perspective. The model assumes the point of view of the organization and it neglects the evaluation needs of all the other stakeholders involved in the training process (Guerci, Bartezzaghi & Solari, 2010). Kaufman and Keller (1994) have suggested that Kirkpatrick's four levels are also incomplete and lead to a too narrowly focus on the evaluation of training alone (Watkins, Leigh, Foshay & Kaufman, 1998). The evaluation framework proposed by Kaufman and Keller (1994) incorporates aspects of

program evaluation, keeps the distinctive four-level features and suggests a five-level evaluation framework. That is, the application of the four levels of training evaluation is

expanded in order to consider the internal and external consequences of all interventions related to performance and organizational improvement. According to these authors, Kirkpatrick's four-level evaluation framework devalues the evaluation of societal impact or the usefulness and availability of organizational resources. They offered four additional aspects (Stokking, 1996, p. 172):

1. Consumer satisfaction and societal contribution as additional evaluation criteria;

2. Evaluation as part of the process of needs assessment and planning;

3. Identification of the desired or expected results and consequences as part of the same process;

4. Availability and quality of resources and efficiency of their use as additional criteria.

Stokking (1996), is equally critical of the Kaufmam and Keller model. Stokking suggests the model lacks clarity in some aspects, such as the distinction between the desired chronology of activities and the aspects of level and importance, or regarding implementation. Implementation and achievement of the learning objectives both integrate Acquisition (Level 2), which should indicate the success of training implementation.

An alternative and widely quoted model is the CIRO (contents/contexts, inputs, reactions and outcomes) model proposed by Warr, Bird and Rackham (1970). The model measures learning/training effectiveness by CIRO elements, both before and after training. The strength of the CIRO model is the measurement of managerial training program and also the effectiveness consideration of objectives (contexts) and training equipment (inputs).

Tzeng and colleagues have suggested that this model does not indicate how measurement takes place and, for this reason, the model does not provide important information regarding the current training situation, which could, certainly, lead to improvements (Tzeng, Chiang & Li, 2007).

The CIPP model (context, input, process and product) proposed by Stufflebeam shares many of the features of CIRO model (Roark, Kim & Mupinga, 2006). However with CIPP, the context provides situational data in order to determine program objectives, input determines the strategies used to achieve the outcomes, product involves program implementation and product involves evaluation of outcomes worth and effectiveness (Khalid, Rehman & Ashraf, 2012).

Bennett (1997) has suggested that the model assumes rationality by decision making and ignores the diversity of interests and multiple interpretations of these agents. Further, Bennett suggests the model is overly abstract and hard to implement in practice.

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

While Kirkpatrick has been the dominant model for organizational evaluation for three decades Phillips' ROI (return-on-investment) framework has emerged in the past decade and has entered the organization evaluation lexicon with its focus on return on investment – a

popular phrase for those conducting investment decisions. The model combines the four levels of evaluation developed by Kirkpatrick and adds a fifth level to measure success in areas of Human Resources function, that is, the ROI measurement compares the monetary benefits from the program with the program costs (Chmielewski & Phillips, 2002).

This evaluation model suggests that while the four factors are useful, without a consideration of the monetary value of specific training initiatives, such as training or coaching, investments should not be considered. The model however has serious limitations, which have largely been ignored in the overt focus on business ROI. One major weakness is the complexity in determining returns on soft aspects of business such as training. In fact, we might suggest such efforts are impossible in non-controlled environments. This is because it is difficult in reality to isolate the effects of the specific intervention, for example training, from other organizational factors which can lead to improvements in performance (Hogan, 2007). These organizational factors can be a change of manager or leadership, to changes to demand for the product or service due to fashion or economic factors, as well as wider impact of other organizational interventions from a pay rise to a change in office layout.

ROI has been used in several training and coaching evaluations with enthusiasm (see for example McGovern et al., 2001). In the McGovern study participants were asked to estimate the value (benefit) of the coaching on key decisions. These estimates where then reduced by 50% and compared with costs. Clearly, no serious scientific study to evaluate the efficacy of a drug or therapy intervention would ask clients to estimate the benefit as part of the evaluation. For this reason alone, this study and others using similar ROI methodology, are in our opinion fundamentally flawed.

Brinkerhoff suggests a six stage approach to evaluation of training that includes the following stages: Goal Setting, Program Design, Program Implementation, Immediate Outcomes, Intermediate or Usage Outcomes, and Impacts and Worth (Kumpikaitė, 2007). Brinkerhoff's model (1989) adds two preliminary levels to Kirkpatrick's model, in order to provide formative evaluation of training needs and the training design (Holton & Naquin, 2005).

This model presents some limitations, since it consists of both formative and summative evaluation, which is only possible in ideal cases where the employer and the training organizers are closely related, where an evaluation design has already been built during the training process, or where there are no competing deadlines or reduced budgets (Holton & Naquin, 2005).

Bushnell (1990) described the IPO (Inputs, Process, Outputs/Outcomes) Model that interprets the evaluation process as cyclical. This model first examines input factors that may influence a program's effectiveness (for example, trainees' qualifications, program design, instructors' quality and qualifications, materials quality, facilities, or equipment). After, it analyses process

factors (such as planning, developing or delivery of the training). Finally, the evaluation of results is organized into evaluation of outputs (short-term results) and evaluation of outcomes (long-term results). Outputs include trainees' reactions, performance or improvement, and outcomes focus on business results (Russ-Eft et al., 2008).

Overall, criticisms of the model are based on its lack of information related to program functioning, or to the specific components that affect the results. Then, there is no way to identify at what point the program failed, because no impact is found (Robertson, 2004).

Holton (1996) proposed the HRD Evaluation and Research Model that hypothesized three outcomes levels: learning, individual performance and organization. According to Holton (1996) these levels are influenced by primary (such as ability, motivation and environmental influences) and secondary factors (for example, those that affect motivation to learn).

Later, Holton (2005) recognizes that a full test of initial HRD Evaluation and Research Model is impossible because the majority of the tools to measure the constructs presented in the model did not exist. For these reasons, the author proposed an updated version of the model by delineating specific constructs that should be measured in each of the conceptual categories proposed (Holton 2005). Kirwan and Birchal (2006) also pointed out that this model solely "describes a sequence of influences on outcomes occurring in a single learning experience and does not demonstrate any feedback loops." (p. 257) and it doesn't indicate any interaction between factors of the same type.

Brinkerhoff (2003) developed the Success Case Method (SCM) for evaluation. According to the author, an SCM study can be used to get answers to any, or all, of four basic questions:

1. What is really happening?

2. What results, if any, is the program helping to produce?

3. What is the value of the results?

4. How could the initiative be improved?

The answers to these questions will give information concerning diverse aspects, such as the way a new innovation is being used; the positive outcomes of a new program or change; identification of organizational units that are using new tools and the success achieved as a result of these new methods; estimation of return-on-investment, support to decision making related to the value a specific program is able to produce, taking into account its current level of impact.

The main disadvantage of SCM is that this model requires some level of judgment regarding what trainers identify as critical success factors on the job, because the model may not identify trainees' problems when returning to work (Casey, 2006).

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

More recently, Dessinger and Moseley (2006) developed the Dessinger-Moseley Full-Scope Evaluation Model. This model blends, in an iterative flow, the benefits of performance improvement and evaluation, and it also integrates formative, summative, confirmative, and meta-evaluation. The main purpose of the model is to formulate judgments about the merit and worth of any performance improvement intervention.

Some of the potential weaknesses of the model are noted by the authors themselves. Dessinger and Moseley (2006) refer that "Full-scope evaluation stays around longer than "regular" evaluation and requires long-term support from the organization and all the stakeholders:" (p. 322).

Table 1 lists the evaluation models previously described, as well as their evaluation criteria.

A recent research study by the Chartered Institute of Personnel and Development (CIPD, 2010) explored the way UK organizations measured and reported the contribution of learning to strategic value. The results identified four main approaches to measuring and reporting on value (CIPD, 2010):

1. Learning function efficiency measures

2. Key performance indicators and benchmark measures

3. Return on investment measures

4. Return on expectation measures.

According to the CIPD, effective evaluation is essential to improve the quality of HR practice. However, to achieve an effective evaluation model is essential. Such models provide the opportunity to place learning and development in the centre of the business; and provide clear measures and metrics of the interventions which have been used (CIPD, 2010).

Table 1: **Ten Popular evaluation models and their criteria**

EVALUATION MODELS	EVALUATION CRITERIA
1. Kirkpatrick's Model	1. Reaction 2. Learning 3. Behavior 4. Results
2. Kaufman's and Keller's Model	1. Enabling and Reaction 2. Acquisition 3. Application

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

	4. Organizational Outputs 5. Societal Outcomes
3. CIRO Model	1. Contents/contexts 2. Inputs 3. Reaction 4. Outcomes
4. CIPP Model	1. Context 2. Input 3. Process 4. Product
5. Phillips Five Level ROI	1. Reaction and Planned Action 2. Learning 3. Applied Learning on the Job 4. Business Results 5. Return on Investment
6. Brinkerhoff's Six Stage Model	1. Goal Setting 2. Program Design 3. Program Implementation 4. Immediate Outcomes 5. Intermediate or Usage Outcomes 6. Impacts and Worth
7. IPO Model	1. Inputs 2. Process 3. Outcomes/Outputs

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

8. HRD Evaluation and Research Model	1. Learning 2. Individual Performance 3. Organization
9. Success Case Method	1. Evaluation Focus and Planning 2. Impact Model Creation 3. Administration of a Survey to Gauge Success Rates 4. Conduction of Interviews with Success and Non-success Instances 5. Formulation of Conclusions
10. Dessinger-Moseley Full-Scope	1. Formative Evaluation 2. Summative Evaluation 3. Confirmative Evaluation 4. Meta Evaluation

The limitations of the models, which are often overly complex, suggest that we need a new comprehensive model for evaluation, which is practical and can be implemented within an organization by human resource managers, as well as offering more advanced methods for researchers for higher level analysis. In the remainder of this paper we aim to set out an alternative model which aims to fill the gaps of existing models, as well as integrating their strengths. Further, our aim has been to develop a practical model which can be used to meet the needs of both practitioners and researchers in the evaluation process. In the next section we describe, the SOAP-M (Self, Other, Achievements, Potential, Meta-analysis) Model.

SOAP-M model for training and coaching evaluation

The model that we propose comprises four levels of analysis which could be used for HR interventions such as training or coaching, with a further level available for formal researchers. The five levels are:

Level 1: Self

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

Level 2: Other

Level 3: Achievements

Level 4: Potential

Level 5: Meta-analysis

The model is summarized in diagram 1. In the following section we describe in detail the levels included in the model, emphasizing their objectives, as well as their relevance.

Level 1: Self

Level 1, like Kirkpatrick's model, is based on self-evaluation of the intervention. We believe the more detailed the questions, the more helpful the evaluation is for the consultant / coach. The focus should be on learning rather than enjoyment.

The Self-evaluation process offers benefits to the participant and the organization. If the questionnaire is well designed it can offer a low-cost way of gathering data as well as useful insights. Firstly, by offering the opportunity for learners to have a voice and to also to reflect on their own learning. Further, we believe that self-evaluation provides valuable and almost instant feedback for the coach or trainer at the end of the training session and allows to the trainer / coach the opportunity of adapting their approach to meet the needs of the audience or individual at the next session.
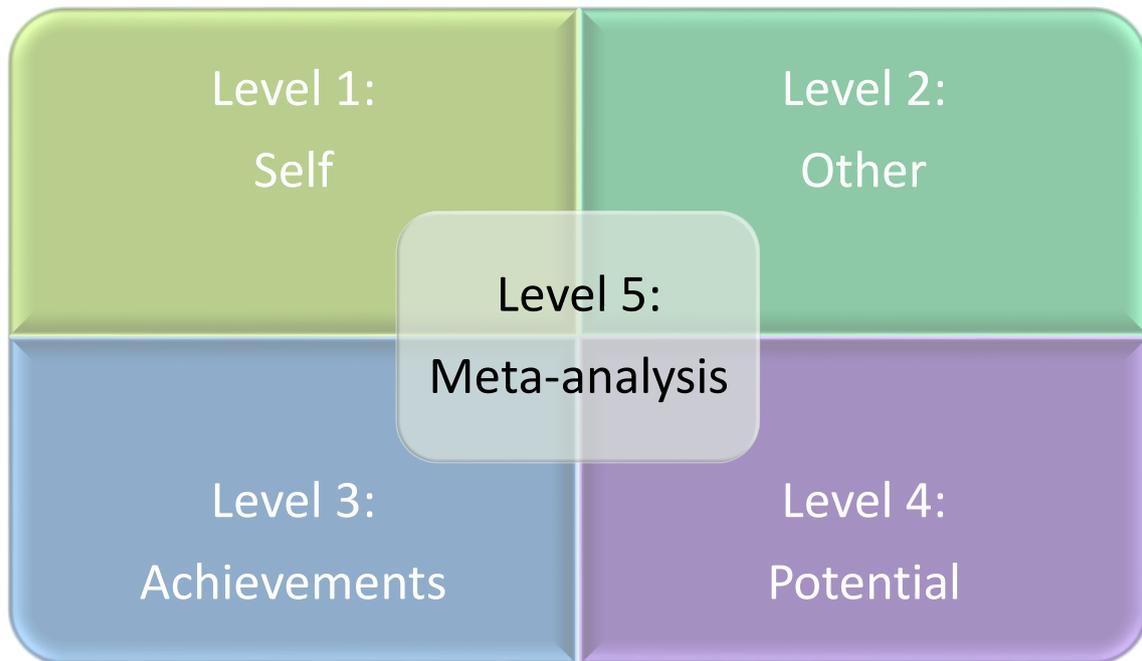
This level might include:

(i)   feedback at the end of an individual intervention or coaching session from coachee or learner on their progress

(ii)   feedback at the end of a coaching assignment / programme from the coachee / learner

An important part of the process, where possible, is the completion of an anonymous feedback form. Where individuals are identified this increases the likelihood that the individual learner's evaluation will be biased.

It could also include the completion of a self-rating competency questionnaire-for example, one which has been designed specifically for the organization or sector. This could be completed as a pre and post intervention questionnaire to measure the individual's self-evaluation of their learning or development from the intervention.

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

Figure 1: SOAP-M Evaluation model

Level 1: Self	Level 2: Other
Level 3: Achievements	Level 4: Potential

Level 5: Meta-analysis

(© Passmore, 2012)

 With the development over the past two decades of increasingly sophisticated psychometric instruments, it could also include the completion of a self-rating psychometric questionnaire, such as the EQi (an emotional intelligence self-evaluation questionnaire – see Perks & Bar-On, 2010) or MTQ48 (a resilience questionnaire – see Clough, Earle & Strycharczyk, 2009). Such instruments allows the self-evaluation results to be compared with a norm group and thus for the individual to receive feedback on how they compare with other managers in the specific domains measured by the questionnaire. With a similar pre and post completion, the added value of the intervention, such as emotional intelligence or resilience training, and the impact on their behavior (as perceived by the participant) can be assessed.

Level 2: Other

At level 2 the evaluation is completed not by the individual themselves but by others. In this case their line manager, peers or stakeholders. Once again this evaluation could use either a predefined framework or model such as an internal (or externally designed) competency questionnaire for example ILM72 (a generic management questionnaire) or a psychometric which allows 360 ratings such as Saville Wave.

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

As with the use of the self rating, using such instruments, the evaluation could use a pre and post assessment by others to evaluate the impact of the training on the individual and their development journey over time.

A number of weaknesses exist at the individual level for evaluation and are inherent in any evaluation method. Firstly, personal factors may account for the change or lack of change. These may include motivation or lack of it as well as ability to learn. Secondly, the use of individual evaluation does not allow for the use of a control group or for statistical testing to provide evidence on the effect size or statistical significance of the result.

These issues can be addressed by the organization in combining individual scores and ratings and using a 'within group' design comparison to produce group results which could be subjected to statistical testing. Such statistical analysis relies on samples of thirty participants of thirty or more and may be more difficult for intervention such as coaching, while it may be relatively easy for training or development centre interventions.

Level 3: Achievements

While level one and two are concerned with behaviour, personality or attitude. Level three shifts the focus to the impact of these newly acquired behaviours on the achievement of key personal tasks. This may include performance against targets set at the annual appraisal or monthly / quarterly goals – such as sales targets or other quantifiable measures. This type of assessment requires SMART goals and some form of comparison with previous goals or with others in the organization, in an attempt to isolate factors such as changing economic conditions or personal circumstances. Level 3 can equally be used to assess organizational level goals such as profit, growth in turn over, market share or stock value /share price, where the responsibility for these rests with the individual, such as a chief executive or sales director.

The higher one moves up the organization, away from individual performance, the harder it is to identify any single intervention which has had an impact. Other variables such as economic conditions, competitor's behaviour and technology change will all play a role. It also becomes more difficult over time as the effect of these contaminating factors grows over time.

Level 4: Potential

Level 4 of the SOAP-M model looks at potential.

Recently developed psychometrics, using high levels of computing power, claim to assess an individual's potential, as well as their actual performance. One example is the Saville Wave. As HR managers know coaching or training interventions may impact on potential as well as short term behaviour – helping individuals to develop thinking skills for example, or emotional skills which they test out in the real world developing emotional maturity as they go. Such aspects are harder to measure through a competence framework and may not be immediately evident in individual achievements, or in self or others assessment of behavior. However, they may be shown up in assessments of the individual's potential.

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

As with previous levels the assessment could take place as pre and post assessment, with completion of the questionnaire at T1 (prior to the intervention) and at T2 (a few weeks or months after the intervention).

We believe the first four levels are all practical methods which most HR professionals could use to evaluate programmes and their impact on individuals as well as groups. As we have identified these work best when there is pre and post assessment providing the opportunity by combining data from individuals to do a 'within' group design analysis.

Level 5: Meta –analysis

Meta-analysis is a more sophisticated data analysis and we believe this is a useful tool for wider reviews of interventions across organizational or sector bodies.

Using individual organizational studies to reveal the impact of an intervention is helpful, but occupational psychology has followed clinical research into placing greater stock on meta-studies. Meta studies combine multiple individual studies looking at the impact of an intervention with many different groups, in different organisations and different cultural context. By grouping studies together the significance of local factors can be reduced and a greater focus to be placed on assessing the intervention.

However while meta analysis has been used extensively in health interventions the number of meta studies is more limited in organization interventions, reflected by the general lack of evidenced based practice within HR / organizational psychology. We hope that this model will encourage researchers to consider meta analysis in training evaluation, as well as for development centres and coaching.

Conclusion

In this paper we have attempted to critically review the ten most popular models offered by writers for use by HR practitioners to help them evaluate HR interventions, such as training or coaching. We note that two models, Kirkpatrick and Phillips models, while popular and well used, have fundamental problems which could be overcome. Other models too have weaknesses. In response, the paper offers a framework which could be used by practitioners to think about and design training evaluations at the start of the training intervention. This model is the SOAP-M model. The model can be used at five levels, self, other, achievement, potential and meta-analysis.

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

References

Alliger, G., & Janak, E. (1989). Kirkpatrick's levels of training criteria: thirty years later. *Personnel Psychology*, *42*, 331-342.

Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning*, 27, 341–347.

Bennett, N. (1997). The Voices of Evaluation. *The Journal of Continuing Education in the Health Professions*, *17*, 198–206.

Bold, E. (2011). Instruments and techniques used in the design and implementation of change management. *Journal of Advanced Research in Management*, *1*(3), 5-13.

Brinkerhoff, R. (2003). *The success case method*. San Francisco: Berrett-Koehler.

Bushnell, D. (1990). Input, process, output: A model for evaluating training. *Training and Development Journal, 44*(3), 41-43.

Casey, M. (2006). *Problem-based Inquiry: An Experiential Approach to Training Evaluation.* University of Akron (unpublished doctoral dissertation).

Chartered Institute of Personnel and Development. (2010) *Evaluating learning and talent development.* London: CIPD. Available at: http://www.cipd.co.uk/hr-resources/factsheets/evaluating-learning-talent-development.aspx

Chmielewski, T, & Phillips, J. (2002). Measuring Return-on-Investment in Government: Issues and Procedures. *Personnel Management*, *31*(2), 255-237.

Clough, P., Earle, K. & Strycharczyk, D. (2009). In J. Passmore (ed.) *Psychometrics in coaching*. London: Kogan Page.

Dessinger, J.C. & Moseley, J.L. (2006). The full scoop on full-scope evaluation. *In* J.A. Pershing (Ed.), *Handbook of Human Performance Technology: Principles, practices, potential*, (pp.312-330). San Francisco: Pfeiffer.

Guerci,M., Bartezzaghi, E, & Solari, L. (2010). Training evaluation in Italian corporate universities: a stakeholder-based analysis. *International Journal of Training and Development 14*(4), 291-308.

Khalid, M., Rehman, C., & Ashraf, M. (2012). Exploring the link between Kirkpatrick (KP) and context, input, process and product (CIPP) training evaluation models, and its effect on training evaluation in public organizations of Pakistan. *African Journal of Business Management, 6*(1), 274-279

Kaufman, R., & Keller, J. (1994). Levels of Evaluation**:** Beyond Kirkpatrick. *Human Resource Development Quarterly*, *5*(4), 371-380.

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

Kaufman, R., Keller, J., & Watkins, R. (1995). What works and what doesn't: Evaluation beyond Kirkpatrick. *P&I*, *35*(2), 8-12.

Kirkpatrick, D. (1979). Techniques for evaluating training programs. *Training and Development Journal*, 78-92.

Kirwan, C., & Birchall, D. (2006). Transfer of learning from management development programmes: testing the Holton model. *International Journal of Training and Development*, *10*(4), 252-268.

Kumpikaitė, V. (2007). Human Resource Training Evaluation. *Engineering Economics*, *5*(55), 30-36.

Hogan (2007). The historical development of program evaluation: exploring the past and present. *Online Journal of Workforce Education and Development*, *II*(4), 1-14.

Holton, E. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, *7*(1), 5-21.

Holton, E. (2005). Holton's Evaluation Model: New Evidence and Construct Elaborations. *Advances in Developing Human Resources, 7*(1), 37-54.

Holton, E., & Naquin, S. (2005). A Critical Analysis of HRD Evaluation Models from a Decision-Making Perspective. *Human Resource Development Quarterly*, *16*(2), 257-280.

McGovern, M., Lindemann, M., Vergara, M., Murphy,S., Barker, L., & Warrenfeltz, R. (2001). Maximizing the Impact of Executive Coaching:Behavioral Change, Organizational Outcomes, and Return on Investment. Manchester Review. 6(1). 1-7. Retrieved on 29 January 2012 from: http://www.coachfederation.org/includes/docs/049ManchesterReviewMaximizingImpactofExecCoaching2.pdf

Passmore, J. (2012). Keynote conference Paper - *Developing an evaluating a coaching culture*, 26 March 2012, Johannesberg, South Africa.

Pellicer, E., Yepes, V., & Rojas, R. (2010). Innovation and Competitiveness in Construction Companies. A Case Study. *Journal of Management Research*, *10*(2), 103-115.

Perks, J. & Bar-On, R. (2009). Coaching for emotionally intelligent leadership, pp55-74. In J. Passmore (ed.) *Leadership in coaching: Working with leaders to develop elite performance*. London: Kogan Page.

Roark, S., Kim, M., & Mupinga, M. (2006). An Exploratory Study of the Extent to Which Medium-Sized Organizations Evaluate Training Programs. *Journal of Business and Training Education*, *15*, 15-20.

Robertson, M. (2004). Building Program Theory for Evaluation: the Process Through a Political Lens. University of Georgia (unpublished doctoral dissertation).

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.

Russ-Eft, D., Bober, M., Teja, I., Foxon, M, & Koszalka, T. (2008). *Evaluator Competencies: Standards for the Practice of Evaluation in Organizations* (pp. 3-12). San Francisco: Jon Wiley & Sons.

Stokking, K. (1998). Levels of Evaluation: Kirkpatrick, Kaufman and Keller, and Beyond. *Human Resource Development Quarterly*, *7*(2), 179-183.

Tzeng, G., Chiang, C., & Li, C. (2007). Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL. *Expert Systems with Applications*, *32*, 1028–1044.

Warr P, Bird M, Rackham N. (1970). *Evaluation of management training: A practical framework, with cases, for evaluating training needs and results.* London: Gower Press.

Watkins, R., Leigh, D., Foshay, R., & Kaufman, R. (1998). Kirkpatrick Plus: Evaluation and Continuous Improvement with a Community Focus. *ETR&D*, *46*(4), 90-96.

Passmore, J. & Velez, M.J (2012). SOAP-M: A training evaluation model for H.R. *Industrial & Commercial Training. 44*(6), 315-325.