PalArch's Journal of Archaeology of Egypt / Egyptology

"A STUDY ON EVALUATION OF TRAINING AND DEVELOPMENT PRACTICES: A THEORETICAL REVIEW"

[*1]Mrs.Vasantha.H , [2]Dr S. Jayanna

[1]Research Scholar, Dept. of Studies in Business Administration, Vijayanagara Sri Krishnadevaraya University, Ballari

[2]Retired Professor, Dept. of Studies in Business Administration, Vijayanagara Sri Krishnadevaraya University, Ballari

ABSTRACT

The success of the training-on which the success of the organizations depends - has to be measured which provides an opportunity for improvement, with this concept the theory of Training Evaluation evolved. Evaluation of training and development means assessment of the impact of training on trainee's performance and behaviour. The paper aims at identifying and studying the theoretical works that have been done so far in the area of Evaluation of Training and Development Practices. It also reviews evaluation studies published and further discusses general evaluation theories in terms of value, use, and evaluator role. The comparison of this literature suggests that evaluation in T&D has been limited by narrow perspectives. The paper concludes by identifying gaps based on this comparison and suggests how evaluation theory may inform future directions of evaluation in human resource development.

Introduction:

Evaluation of training and development is the most essential aspect of training programme. Generally all good training and development programmes start with identification of training and development needs and ends with evaluation of training (Gopal, 2009). Training evaluation ensures that whether candidates are able to implement their learning in their respective work place or to the regular routines (Nagar, 2009). Phillips (1991) defined evaluation as a systematic process to determine the worth, value or meaning of something. Evaluation of training and development involves assessing whether it is achieving its objectives, it is effective or not.Given the importance of evaluation as the last component in the training process and as a key opportunity for training improvement, training evaluation theory has been developed over the last fifty years (Phan, 2008). The pioneer to

introduce training evaluation theory is Donald Kirkpatrick, who in 1959 proposed a simple four-levels of evaluation of training programme (Tamkin et al., 2002; Bassi and Cheney, 1997; Bernthal, 1995). In 1994, Donald Kirkpatrick's book "Evaluating Training Programs" defined his originally published ideas of 1959, thereby further increasing awareness of them, so that his theory has now become arguably the most widely used and popular model for the evaluation of training and learning. Despite the model's advantages of simplicity and practicability, since the outset, it has received several fundamental and recently increasing criticisms. It resulted in emergence of other evaluation theories likeHolton, Kaufman and Watkins in 1996, Wang,Dou and Li in2002 ,Matson in 2003 and so on.

Objectives:
1.The paper aims at identifying and studying the theoretical works that have been done so far in the area of Evaluation of Training and Development Practices.
2. The study also intends to review evaluation studies published and further discusses general evaluation theories in terms of value, use, and evaluator role.

Research Methodology:
    The present enquiry is conceptual study. The work is based on secondary sources of information and data. The information about the problem is collected from the Research Journals, Trade Magazines, and the Internet. The researcher has attempted a relatively comphrensive survey of theoretical works.

CRITIQUE OF EXISTING THEORIES AND MODELS OF EVALUATION
Kirkpatrick's four-levels of evaluation of training programmeare now considered an industry standard across the HR and training communities(www.businessballs.com). Subsequently, the Kirkpatrick's framework for evaluation of training, also known as the four-level (reactions – a measure of satisfaction, learning – a measure of learning, behavior – a measure of behavior change, results- a measure of results) evaluation model, is acknowledged by many practitioners as the standard framework in the field (Stone and Watson, 1999) (Phillips, 1997); (Kirkpatrick, 2009). These four levels cover much of the criteria for measuring success. This framework promoted the awareness of the importance of thinking about and assessing training in business terms (Wang, 2003). Kirkpatrick's classic work is still considered as the foundation for training evaluation with the advantages of its practical and simple nature (Tamkin et al. (2002), Alliger and Janak (1989), Aragon-Sanchez et al. (2003), and Tennant et al. 2002). Due to its straightforwardnature, the four-level model has become the standard in the field of evaluation(Holton, 1996; Bramley and Kitson, 1994) and the most popularly applied (Phillips, 1991; Donovan et al., 2001). Besides the simplicity and practicality, the model'sstrength was also recognized in its focus on the changes in behavioral outcomes ofthe learners involved in the training (Tennant et al., 2002

Holton (1996)argued Kirkpatrick's four-level model of evaluation is flawed and is better described as a taxonomy. He discussed weak relationships among the levels and how the model fails to account for the complex system of influences on training outcomes that exists in organizations. Holton (1996) proposed a significantly

different model by removing reaction measures, focusing on individual performance instead of behavior, and including intervening variables. This model includes links among learning, individual performance, and organizational performance with intervening variables such as ability, environmental, and motivation elements as well as secondary influences such as personality and attitudes. Holton acknowledged his proposed model is complex and requires a significant investment to collect all the data. However, he suggested a key to improve evaluation is to stop asking "happiness" questions and focus on self-reported estimates on the impact to individual and organization performance.

Wang, Dou, and Li (2002) introduced a model for measuring return on investment in a systems approach. They defined return on investment in HRD as "any economic return, monetary or non-monetary that accrue through HRD investments (p. 212)." They noted a primary reason that returns on investment measurement of HRD interventions is still in its infancy is due to dynamic human behaviours not found in fixed assets. From this definition they introduced a systems approach for measuring HRD interventions. They developed a model that identifies HRD as an organizational subsystem impacting business outcomes and introduced in-depth mathematical calculations to determine the "fair share" of HRD intervention impacts on firm performance. While this model provides a new way of examining the returns of HRD investments it is tremendously complex to implement.

Kaufman and Watkins (1996) noted the complexity of return-on-investment analysis is often more detailed than an organization desires or is able to measure. Using the Organization Elements Model (OEM) as their foundation, the authors established the costs-consequence analysis as a "coarse-grained estimate" of what one puts into a system and what one gets out of the system. Kaufman and Watkins provided an in-depth discussion of the elements and ways to measure costs and consequences for each element. For example, auditing and accounting can measure inputs, cost effectiveness can measure products, and cost-utility can measure outcomes of the OEM. This proposed costs-consequence analysis provides an alternative to return on investment.

While return on investments and related financial measures of training impact have been an area of research emphasis, Pershing and Pershing (2001) reminded us that the most common evaluation approach remains reaction measures or level one of Kirkpatrick's model. Based on the high frequency of reaction evaluation usage relative to usage of other levels of evaluation, they investigated design and implementation elements of this level. Through their analysis, the authors concluded that if the most prevalent level of evaluation is to become more useful for evaluators, instructors, and learners then design and content need to adhere to uniform and consistent standards.

While the previously discussed research critiqued established models and proposed new models for HRD evaluation, a second set of literature in HRDQ has emerged in recent years. This literature establishes a trend that moves away from criticisms of current evaluation models and towards understanding the relationship of evaluation and organizational decision-making processes.

Bober and Bartlett (2004) noted that evaluators need to focus on the utilization of results and not just the methods and processes for collecting data. They explored

how various organization members use evaluation data and for what intended purposes. The authors identified that evaluation data can be used for numerous purposes expressed as a continuum from direct and short-term uses to conceptual uses oriented towards future policy-based decisions. This continuum represents actions that can be derived from evaluation such as improvement of training, accountability, and program justification and continuation. They identified three primary users of evaluation results; evaluation staff, instructional designers, and upper-level management. They further identified factors influencing the usage of evaluation data that include communication quality, timeliness, quality, credibility, and commitment to evaluation. This study emphasized knowing stakeholder needs and their planned use of evaluation data.

Mattson (2003) examined the effects that alternative evaluation reports have on managerial support for HRD interventions. The author utilized three report methods (utility analysis, COT report, and anecdotal evaluation) to ascertain whether the type of evaluation report accounted for significant differences in perceived usefulness. Mattson identified that evaluation usefulness is linked with credibility of the information source mediated by expertise and trustworthiness. Mattson concluded that managers prefer financial results information over anecdotal information. In fact, anecdotal reporting was identified as the least useful for management. Furthermore, he noted that management responded more favourably to business results evaluation as opposed to reaction evaluations.

Holton and Naquin (2005) noted that all evaluation models have decision-making as their core output. The primary purpose of evaluation in organizations is to contribute to better decision-making. They argued that evaluation models are not good decision-making models. The authors explored the literature of decision-making theories and the role of bounded rationality. They concluded that while evaluation models can be categorized as rational-economic decision-making, researchers have long recognized that people do not follow rational-economic patterns in their decision-making. They identified a gap for the field to further explore in order to discover new evaluation models that follow more naturalistic decision-making behaviors, which identify economic factors as only one of several factors to evaluate HRD interventions.

As the field has transitioned from critiques and adaptations of Kirkpatrick's model to connecting evaluation with organizational decision-making, two important issues for research emerge. First, there remains a missing linkage among the four levels of Kirkpatrick's model. This is evident by the extensive use of reaction evaluation and minimal use other levels of evaluation. Second, while practitioners are interested in financial data to evaluate the impact of human resource development investments, economic factors may represent only a subset of several factors to consider. Cascio (2003) pointed out that proper framing of the message is essential and necessary to articulate the links between HRD interventions and organizational success. Holton and Naquin (2005) suggested evaluation models in HRD do not assist decision-making because of the failure to follow more naturalistic decision-making approaches that may offer a proper framing of these messages. Finally, Torres (2004) emphasized that as the field moves forward in evaluation research it is essential to understand the role of context in evaluation.

To further address the issues that emerge in the review of HRD literature, it is

necessary to examine broad evaluation theories. This review of theoretical approaches to evaluation outside the field of HRD can provide insight into new research directions and guide a maximum understanding of evaluation (Torraco, 1997).

Informing through Evaluation Theories

While the practice of evaluation has been a priority of many fields including HRD, evaluation theory has been less focused and even neglected in the implementation of the evaluation practice (Shadish, Cook, &Levition, 1991). However, comprehension of theory should precede application of evaluation practice, because theory can guide the integrity and wholeness of applied efforts such as evaluation (Swanson & Holton, 2001). In other words, "evaluation theory tells us when, where, and why some methods should be applied and others not (Shadish, et. al., 1991, p. 34)." Since evaluation theories have varied, it is necessary to understand the various perspectives in order to grasp a full appreciation of the theory that informs this area of inquiry.

To inform the current status of HRD evaluation research, the authors discuss various evaluation theorists' perspectives based on **three components, value, use, and role,** which ground evaluation theories (Alkin& Christie, 2004; Shadish, et. al., 1991).

Value

The ontological and epistemological assumptions and values in evaluations provide fundamental logic for guiding the use, the method, as well as the role of evaluators. Objective-oriented evaluation, which emphasizes specification of objectives and measurement of outcomes, is strongly rooted in behaviorism of the early twentieth century. Tyler (1942) believed that evaluation objectives are measurable and that there are objective standards and norms to be achieved for making judgments. In his notion, evaluators' and stakeholders' values can be thoroughly controlled. However, other scholars recognized that values influenced evaluation and provided multiple perspectives (Campbell, 1984; Shadish&Luellen, 2004; Cook, 2004). While Campbell (1984) thought that stakeholders' and evaluators' value could be minimized, Cook (2004) viewed stakeholders' values as essential input in deciding evaluation objectives and considered this variation of values in the evaluation context.

The notion of value-free evaluation was fully challenged by the next generation of theorists (Boruch, 2004; Rossi, 2004; Chen, 2005; Cronbach, 1980; Scriven, 1993; Stake, 2003; Stufflebeam, 2003; Preskill, 1998). These scholars generally maintained that multiple realities and various perspectives should be considered in evaluation practices. However, their thoughts differed from each other in terms of how they perceive stakeholders' perspectives and whose values are ultimately considered in their evaluation practices.

One group of evaluation scholars advocated a *theory-driven evaluation*. This approach emphasized acknowledging stakeholders' prescriptive and descriptive assumptions when building a program theory (Boruch, 2004; Rossi, 2004; Chen, 2005). This view is not interested in building a grand and general theory rather a local level of plausible and defensible theory that could be used in evaluating

programs based on the local context.

Other scholars agreed with the theory-driven evaluation value in that stakeholders' perspectives are important to consider in the evaluation process (Stake, 2003, 2004; Stufflebeam, 2003, 2004; Cousins, 2003, 2004) However, while theory-oriented scholars value stakeholders' perspectives for building program logic, Stake (2003, 2004) viewed stakeholders' perspectives and values as epistemological sources of evaluation for evaluators' understanding. That is, Stake ultimately championed evaluators' expertise rather than stakeholders' program logic in decision making. This value differed from Stufflebeam (2003, 2004) who advocated *decision-oriented evaluation.* Stufflebeam thought stakeholders should be guided by information for their judgments and decision making.

Cousins (2003, 2004) advocated a pragmatic *utilization focused evaluation*, and Preskill (1998, 2004) advocated *participatory evaluation* in terms of empowerment. Stufflebeam leaned towards evaluator provided information for decision making, while Cousins and Preskill preferred to value program practitioners' intimate knowledge and perspectives about the program and its context. Preskill placed additional value on stakeholders' empowerment through organizational learning than evaluator-directed approaches.

Other theorists questioned whose value should ultimately be considered in the evaluation process. For example, Cronbach (1980) advocated a *policy-oriented evaluation* and described that the benefit of the larger community could be reconciled with commitments to a sponsor. Similarly, Weiss (2004) advocated policy-oriented evaluation, but emphasized program recipients who tend to be the most deprived groups relative to the evaluation sponsors. Scriven's (2004) *consumer-oriented evaluation* corresponded with Weiss by valuing the recipients of the program services which became the base of a goal free evaluation.

Use

Two main theoretical approaches of evaluation literature can be identified in understanding the use of evaluation. One paradigm, the instrumental view, prioritizes the importance of the process and unbiased results. The second paradigm is concerned with the usefulness of the data. These two approaches coincide with the literature review of HRDQ articles that emphasized the process and results as well as examined the exploration of the relationship between evaluation and organizational decision-making.

The instrumental view of evaluation seeks to identify program effectiveness and improvement (Tyler, 1942; Campbell, 1981; Scriven, 1993). Even though Campbell (1981) identified evaluation as the experimenting society for policy-makers, his focus has been on providing unbiased evaluation results rather than ensuring the use of evaluation results. This view of evaluation does not prioritize facilitating and ensuring the use of evaluation. In other words, the primary concern of this evaluation approach is not ensuring evaluation results are being used, rather getting the process of evaluation right (Rossi, 2004). This view does generally consider stakeholders' involvement in evaluation as an approach to increase the likelihood of the use of evaluation results.

The second view of evaluation use is concerned with how the evaluation

results could be used. Stufflebeam (2003) argued that evaluation should be used from the start to the end of a program. Stufflebeam's (2003) CIPP model provided guidance for formative and summative evaluation and ensuring a timely utilization of evaluation information. Cronbach (1980) stressed that evaluation contributes to policy-shaping and that the use of evaluation is subject to the function of the political system. This provides a holistic perspective in understanding the use of evaluation by recognizing the political system and context. Cronbach (1980) further stressed that communication with the policy-shaping community could facilitate the use of evaluation results. While Stufflebeam's view of use involved a timely, specific, and local level of use, Cronbach's view provided a broader notion of evaluation use.

Stake (2003) agreed with Cronbach's view that evaluators do not have control over the use of evaluation. Stake thought that evaluation is not for the evaluators' instrumental use, but for the fairly conceptual and symbolic use for stakeholders' future use since stakeholders determine the quality of program.

Cousins (2003) introduced a more active and comprehensive view of evaluation use by focusing on the process of evaluation. Cousins stressed responsiveness to stakeholders' needs. This explicit recognition of responsiveness differed from Stake's (2003, 2004) responsive evaluation. Stake's approach is epistemological-oriented and responsive to activity and program. On the contrary, Cousins (2003, 2004) view is pragmatic utilization-oriented through perceived evaluation as a responsive approach to the stakeholders' needs. While Stake does not allow stakeholders' participation in evaluation, Cousins argued that stakeholders should learn and develop their evaluation research skills through participation in the evaluation process for self-critique and self-determination, through which their evaluation will be actively used.

Similar to Cousins, Preskill's (1998) empowerment and participatory evaluation is also strongly oriented to the process use. Both scholars stressed stakeholders' individual and organizational learning for their self-development. Moreover, Cousins and Preskill have the same rationale stakeholder participation and empowerment though organizational learning can be justified by the enhanced use of evaluation. However, Preskill's emphasis of stakeholders' learning in evaluation is fairly different from Cronbach's educative evaluation in that while Cronbach's educative thoughts are toward identifying issues related to policy making, Preskill's participatory evaluation is ultimately intended for stakeholders' empowerment.

Role

Perspectives of the evaluator's role has expanded from narrow notions of program outcome reporting to views of more engaged and directive roles even to the level of promoting change through evaluation. In the early traditions of evaluation, evaluator roles were identified simply as scientists and measurement developers. An evaluator represented a co-developer of program objectives with expertise in scientific testing and measurement (Tyler, 1942). Furthermore, an evaluator's role was to strictly provide information and advice in a decision-making process (Campbell, 1981, 1984).

Cronbach (1980) broadened this narrow definition of the evaluator's role to describe evaluators as educators who illuminate and stimulate decision-making by

providing constructive feedback to clients in order to facilitate enlightenment. While this perspective did not suggest that evaluators make decisions, they did perceive that evaluators are to influence decision and policy making processes through timely information (Cronbach, 1980; Stufflebeam, 2003, 2004). Stake (2003, 2004) also recognized the long-term educative role of an evaluator. However, he described an evaluator's role in even more engaged terms as an expert who identifies, negotiates, and ultimately selects a few issues among all voices to interpret the quality of a program. This, in turn, would enhance stakeholders' learning and application of evaluation to other situations in the future. In other words, Stake viewed an evaluator as an ethnographer whose role is to hear the participants' voices and concerns.

In general, these identified roles of evaluators coincide with the scholar's views of the extent to which stakeholders should participate in the evaluation process. Early traditions of evaluation did not consider stakeholders participation necessary and viewed evaluators as those who are supposed to discover and know everything (Madaus, 2004). Other perspectives emphasized collaboration through stakeholders' participation, to the point of identifying stakeholders' assumptions and mining their knowledge from the evaluators' perspective. And other scholars viewed stakeholders as experts who provide program logic and context so that they can make informed judgments. From this view, the meaningful participation of stakeholders is compulsory. Therefore, the evaluator's role is to enhance stakeholders' meaningful participation. Hence, team building, facilitation, negotiation, conflict resolution, and interpersonal skills are requisite skills of an evaluator.

Conclusion:
Training evaluation is the most important aspect of training and development. All models are descriptive and subjective in nature. From the above discussion it has been found out that Kirkpatrick model is widely used model, however there is a wide range of training evaluation models available, each with their own unique advantages and disadvantages. One crucial point is that none of these models need to be implemented in their entirety; they can always be adapted to suit the needs and requirements of a specific organization. With careful consideration, an organization can find the most suitable approach for their situation and circumstances.The various perspectives of the value, use, and role of evaluation provide a broad insight into the processes and intended outcomes of evaluation.This review brings to light the assumptions that underlie the implementation of evaluations. Understanding these insights in the context of HRD evaluation traditions and the resultant gaps in our field's knowledge, provide some directions for future HRD research and practice.

References
1.Alkin, M. C. (2004). Evaluation roots: Tracing theorists' views and influence. Thousand Oaks, CA: Sage.
2.Alvarez, K., Salas, E., &Garofano, C. M. (2004). An integrated model of training evaluation and effectiveness. Human Resource Development Review, 3(4), 385-416.
3.Bober, C. & Bartlett, K. (2004). The utilization of training program evaluation in corporate universities. Human Resource Development Quarterly, 15(4), 363-383.
4.Boruch, R. F. (2004). A trialist's notes on evaluation theory and roots. In M. C.

Alkin (Ed.) Evaluation roots: tracing theorists' views and influences (pp. 114-121). Thousand Oaks, CA: Sage.

5. Campbell, D.T. (1984). Can we be scientific in applied social science? In R. F. Conner, D. G. Altman, & C. Jackson, Evaluation studies: Review annual (Vol. 9). Beverly Hills, CA: Sage.

6. Cascio, W. (2003). Invited reaction: the effects of alternative reports of human resource development results on managerial support. Human Resource Development Quarterly, 14(2), 153-158.

7. Chen, H.T. (2005). Practical program evaluation: Assessing and improving planning, implementation, and effectiveness. Thousand Oaks, CA: Sage.

8. Collins, D. (2002). Performance-level evaluation methods used in management delopment studies from 1986 to 2000. Human Resource Development Review, 1 (1), 91-110.

9. Cook, T. D. (2004). Causal generalization: How Campbell and Cronbach influenced my theoretical thinking on this topic, including in Shadish, Cook, and Campbell. In M. C. Alkin (Ed.) Evaluation roots: Tracing theorists' views and influences (pp. 88-113). Thousand Oaks, CA: Sage.

10. 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.

11. Kaufman, R. & Watkins, R. (1996). Costs-consequences analysis. Human Resource Development Quarterly, 7(1), 87-100.

12. Mattson, B. (2003). The effects of alternative reports on human resource development results on managerial support. Human Resource Development Quarterly, 14(2), 127-151.

13. Rossi, P. H. (2004). My views of evaluation and their origins. In M. C. Alkin (Ed.) Evaluation roots: Tracing theorists' views and influences (pp. 122-131). Thousand Oaks, CA: Sage.

14. Scriven, M. (1993). Nature of evaluation. In M. Scriven (Ed.), New directions for program evaluation: Hard-won lessons in program evaluation (pp. 5-48).Thousand Oaks, CA: Sage.

15. W. R. &Luellen, J. K. (2004). Donald Campbell: The accidental evaluator. In M. C. Alkin (Ed.) Evaluation roots: Tracing theorists' views and influences (pp. 80-87). Thousand Oaks, CA: Sage. 12

16. Stufflebeam, D. L. (2003). The CIPP model for evaluation. In T. Kellaghan, & D.L. Stufflebeam (Eds.) International handbook of educational evaluation (pp 31-62). Dordrecht: Kluwer Academic Publishers.

17. Torraco, R.J. (1997). Theory-building research methods, In R.A. Swanson & E.F. Holton (Eds.), Human resource development research handbook, linking research and practice (pp 114-137). San Francisco: Berrett-Koehler.

18. Wang, G. (2003). On the two-way customer service model in HRD practice. Human Resource Development Review, 2(4), 453-458.