UNIVERSITI TEKNOLOGI MARA

INSPIRED 2024
IPOH INTERNATIONAL SUMMIT ON PROFESSIONALISM, RESEARCH & EDUCATION

e-Proceedings

2024 ICOAB
INTERNATIONAL CONFERENCE ON ACCOUNTING & BUSINESS

Ethics and Integrity in Accounting and Business: Building Trust in an Intricate World

ORGANISED BY:

FACULTY OF ACCOUNTANCY, UiTM PERAK BRANCH
TAPAH CAMPUS

In Collaboration With

UNIVERSITI TEKNOLOGI MARA | Institut Penyelidikan Perakaunan
UNIVERSITAS INDONESIA — Veritas, Probitas, Iustitia
UNIVERSITAS ISLAM INDONESIA
universitas paramadina
UTama UNIVERSITAS WIDYATAMA

PROCEEDINGS OF THE 2ND INTERNATIONAL CONFERENCE ON ACCOUNTING AND BUSINESS

(ICAB2024)

Ethics and Integrity in Accounting and Business: Building Trust in an Intricate World

e ISBN 978-967-2776-35-2

FACULTY OF ACCOUNTANCY
UNIVERSITI TEKNOLOGI MARA, PERAK BRANCH, TAPAH CAMPUS, MALAYSIA

Perpustakaan Negara Malaysia Cataloguing in Publication Data

No e-ISBN: 978-967-2776-35-2

Cover Design: Graphic Design and Web Page Team
Typesetting : ICAB 2024 Proceeding Team

Evaluating Training Effectiveness Using Kirkpatrick's Model: A Conceptual Paper

Agung Wahyu Setyadi*, Didin Hikmah Perkasa, Ahmad Azmy, Agung Surya Dwianto, Lingga Yuliana
Paramadina University , Mampang Prpt., Mampang Prpt. District, South Jakarta City, Special Capital Region of Jakarta 12790, Indonesia

*Corresponding email: agung.setyadi@students.paramadina.ac.id

Abstract

Evaluating training is a vital aspect of professional growth, empowering employees to refine their competencies, unlock potential, and adjust work behaviors, thereby enhancing their knowledge, skills, abilities, and motivation, leading to improved work performance. This paper explores Kirkpatrick's Four-Level Model and its effectiveness in employee training. Drawing from theories in training evaluation, including program evaluation and research evaluation, the paper underscores the enduring relevance of the original Kirkpatrick model while advocating for the adoption of the World Kirkpatrick Four-Level Model to enhance evaluation comprehensiveness. Additionally, recognizing the Reductionist nature of the Kirkpatrick model, the paper suggests integrating Complexity Theory and Systems Theory to fortify its theoretical underpinnings and address its limitations.

Keywords
Training, Effectivesness, Training Evaluation, Kirkpatrick's Model

1.0 Introduction

In the contemporary era of Industry 4.0, the competitiveness within the labor market has intensified significantly. The demand for high-quality, competent, and loyal human resources (HR) is crucial for organizational success, as their performance critically impacts organizational productivity (Priyono & Perkasa, 2024). Training is essential to refresh and enhance the knowledge, abilities, and outlook of employees. It is a systematic process aimed at shaping employees' competencies, developing their potential, and altering work behavior, including enhancing their knowledge, abilities, skills, and motivation (Demerouti & Peeters, 2018). It also has transformative impact on driving the improvement of their work performances (Gil et al., 2023). Without adequate training, employees' full potential remains untapped (Giday et al., 2023). Conversely, well-trained staff can help achieve organizational objectives. They become more productive and are better equipped to handle complex tasks or perform their existing responsibilities more effectively and efficiently (Pedrini & Cappiello, 2022).

Consequently, in the field of organisational development and human resource management, the evaluation of the effectiveness of training is of critical importance in order to ensure that investments in employee development yield optimal returns. It indicates that effective training brings benefits not only to the company but also to the participants (Asghar et al., 2022). Urbancová et al. (2021) argued that effective training evaluation serves as a tool for the development and success of the entire organization and for the enhancement of its business performance. Furthermore, the evaluation enables the ongoing improvement of training quality and the identification of organizational needs (Bęś & Strzałkowski, 2024).

While the majority of formal training and development programs undergo evaluation, significant limitations exist. The predominant method of evaluation involves assessing employee responses to the

training, typically conducted on the final day, with the employee responses being inadequate (Steensma & Groeneveld, 2010). Additionally, the challenges in training evaluation are also about the efficient budget to have effective results on the training. Furthermore, the inadequate and poorly executed appraisal system, coupled with a lack of awareness of or access to evaluation methods and tools, may impede the effectiveness of training (Saad & Mat, 2013; Alsalamah & Callinan, 2021). As a result, a number of models have been proposed for the purpose of evaluating the impact of training program. One of the most renowned and extensively applied evaluation models for training and development initiatives is the Four-level Evaluation Model, which was introduced by Donald Kirkpatrick in 1959. The Kirkpatrick Four-level model consists of four levels of training: (1) Reactions; (2) Learning; (3) Behavior; and (4) Results. This model has undergone rigorous scrutiny and has gained increasing recognition, emerging as one of the most embraced and impactful frameworks in the field. Yet, it established a systematic structure for evaluating outcomes and effects, considering both individual and organizational performance dimensions. Kirkpatrick recognized the necessity of creating an analytical framework to evaluate the degree to which training initiatives enhance learners' acquisition of job-related skills, improve service quality and safety, and impact the community. Some studies have empirically proven the effectiveness of this model in different industries (Moldovan, 2016; Alsalamah & Callinan, 2021; El Nsouli et al., 2023). Steensma & Groeneveld (2010), for example, suggested that Kirkpatrick's Four-Level Model is sufficient to be utilized for evaluating the effectiveness of management training. However, it is also important to further synthesize concepts related to Training Evaluation Effectiveness that can enrich the understanding of the literature for further research in the future, given the shortcomings identified in the adoption Kirkpatrick's Four-Level Model (Thalheimer, 2018).

Zhang et al. (2019) suggest that further research is needed on employee training in modern companies. Longo et al. (2023) emphasizes that, in the technological era, it is crucial to seek new approaches to enhance training and development effectiveness, as traditional methods are becoming increasingly ineffective due to their inability to keep pace with technological advancements and industry changes (Fraile et al., 2023). Furthermore, Pedrini et al. (2023) argue that research on the training evaluations has been neglected in certain industries, which most of the study focus on empiral evidence. Therefore, this paper aims to explore the training evaluation effectiveness of Kirkpatrick's Four-Level Model, and to seek for a theoretical framework for evaluation effectiveness that supports the model.

2.0 Theoretical Background

2.1 Training Evaluation

Evaluation, in a broader sense, is defined as "an integral part of the design, development, delivery, and implementation of programs" (Phillips & Phillips, 2016). Evaluation is needed in training program, as it typically utilizes objective-oriented and systems-oriented methods to select the most suitable approach for assessing safety training, thus ensuring its efficacy and purposefulness (Asari & Leman, 2015). According to Gebrehiwot and Elantheraiyan (2023), training evaluation entails assessing the training program to ensure its effective implementation and to measure the returns on investment relative to the resources allocated to the operation. Noe (2017, p. 216) refers the term "training evaluation" as "the process of collecting the outcomes needed to determine whether training is effective." Similarly, Brown (2007, p. 820) defines the training evaluation as "a process that may be used to determine the effectiveness and/or efficiency of instructional programmes."

As outlined by Kirkpatrick (1998), the evaluation of training is highly important for the improvement of training programs and for identifying potential issues or weaknesses in the evaluation process. Moreover, evaluation is utilized to confirm the efficiency of training, making it the preferred approach

for this purpose (Khandker et al., 2010). Consequently, training evaluation offers a number of benefits. It enables the identification of the necessity for innovation through the implementation of high-quality, cost-effective, just-in-time training for all employees. Moreover, for employees, different training initiatives are linked to the extent of their impact on both employability and income over the medium to long run (Ghirelli et al., 2019). Similarly, on an organizational level, such evaluation contributes to measuring the effectiveness in order to deliver the world-class training required to maintain a competitive advantage (Kirkpatrick & Kirkpatrick, 2006; Noe, 2017).

2.2 Training Effectiveness

Noe (2017, p. 216) defines training effectiveness as "the benefits that the company and the trainees receive from training. Benefits for trainees may include learning new skills or behavior." Training effectiveness refers to "the extent to which trainees (and their organization) benefit as intended from training (Brown, 2007, p. 820). Thus, Ford (2007, p. 833) emphasized that "training effectiveness evaluates whether or not the training achieved its intended outcomes." The evaluation of training effectiveness guides the ongoing enhancement of its quality, wherein selecting the appropriate training method is customized to suit the participants' group, considering their varying experience, skills, and knowledge levels (Bęś & Strzałkowski, 2024). Beside those attirbutes, training effectiveness also present valuable chances to bolster organizational effectiveness (Al Qudah et al., 2018). Urbancová et al. (2021) argued that the process of evaluating the effectiveness of a training and development initiative allows for the refinement of future alterations, with the aim of enhancing processes and amplifying outcomes across all organisations (e.g., knowledge, skills, attitudes). Therefore, the measure of training's effectiveness should rely on how well participants can put into practice the knowledge, skills, and attitudes they have gained from the training (Brinia & Efstathiou, 2012). The outcome of a training program depends to a considerable extent on the degree of accuracy with which its effectiveness is evaluated. (Karim et al., 2018).

2.3 Kirkpatrick's Four-Level Model

The four level of evaluation is a model represent a sequential method for program evaluation, with each level's significance influencing the subsequent one (Kirkpatrick & Kirkpatrick, 2006). The purpose of this model is to inspire training directors to recognize the importance of evaluation and to enhance their efforts in evaluating their training programs (Johnson & Dick, 2012. p. 99). Kirkpatrick (1998) delineates that evaluation serves specific purposes: to validate the allocation of a training budget by demonstrating the alignment of the training program with organizational goals and objectives, to ascertain the continuation status of a training program, and to gather insights for enhancing future training initiatives. Furthermore, Kirkpatrick and Kirkpatrick (2006) asserted that the models exhibit interconnectedness concerning (1) discovering pathways for continual enhancement, (2) evaluating the effectiveness of endeavors and those involved, and (3) ensuring that organizations' training investments yield a suitable return on investment. This model has shifted the emphasis of training assessment practices towards outcome-oriented methodologies (Bates, 2004; Alsalamah & Callinan, 2021). Figure 1 illustrates a spectrum of evaluation from individual to organizational purposes.

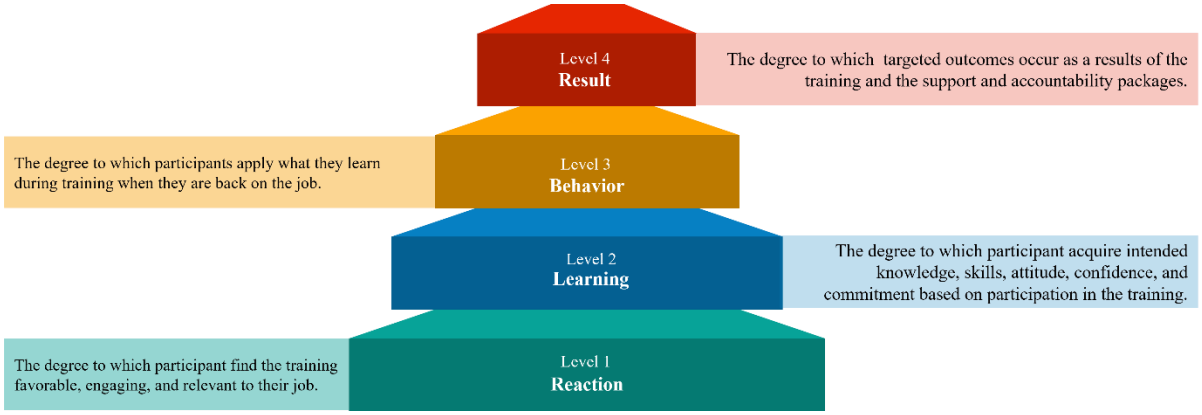

Figure 1: Kirkpatrick's Four Level Evaluation Model
Source: Adapted from Kirkpatrick (1998, 2006)

Each level of evaluation of the model (Kirkpatrick, 1998; Kirkpatrick & Kirkpatrick, 2006, 2015) has different emphases and is summarized as follows:

2.3.1 Reaction Level

Reaction is the first level of Kirkpatrick's Four-Level Model. It measures the extent of trainee reaction to the training program in form of satisfaction or impressions of the program. Evaluation of reaction is important because a trainee's positive or negative reactions towards a training course influence learning. If the trainees give positive feedback or reactions to the training, they are motivated to learn and bring about more learning. An increase in learning leads to improvised knowledge, skills, and a positive attitude towards innovation introduced during the training.

2.3.2 Learning Level

Level 2 is Learning, which measures how much trainees gain knowledge, improve skills and change their attitudes after attending the course. The necessity for evaluating learning arises from the assumption that behavioural changes will not occur unless one or more of the specified learning objectives have been attained. The evaluation of learning is of paramount importance, as it precedes behavioural changes. In the absence of learning, behavioural shifts are unlikely. Nevertheless, there exists a differentiation in measurement that indicates the extent to which learning has occurred. Thus, level 2 evaluation aims to ascertain what knowledge was acquired, what skills were enhanced, and what attitudinal change resulted from the training.

2.3.3 Behaviour Level

Level 3 evaluation aims to measure the extent to which knowledge, skills, and attitudes learned in training are applied in the workplace, essentially measuring the change in job behavior resulting from attending a training program. In other words, this level of evaluation measures the transfer of what has been learned back at the workplace or training transfer. Nonetheless, improvements in employees' skills and knowledge immediately post-training do not guarantee behavioural changes.

2.3.4 Result Level

Level 4 represents the most significant and arguably the most challenging stage of the process, as it seeks to ascertain the extent to which the training program has influenced the final results. The results are determined by the training program's objectives as well as the stakeholders' expectations (who are administratively responsible for the employee being trained). According to Kirkpatrick and Kirkpatrick (2016), results represent a synthesis of the organizational mission and objectives. In a business context

driven by profit, this entails the effective delivery of products or services to the market. In non-profit, governmental, or military contexts, it involves the fulfillment of the designated mission.

3.0 Methods

This paper presents the concept of Kirkpatrick's four-level training evaluation model and aims to explore its effectiveness in evaluating training programs conducted by companies for their employees. A Narrative Literature Review was employed to gather relevant data and articles on the topic. Due to the inherently non-systematic nature of this approach, the Narrative Literature Review does not follow a defined protocol or a specified search strategy, nor does it adhere to formal guidelines, as it is guided solely by a topic of interest (Gregory & Denniss, 2018; Demiris et al., 2019). Consequently, this paper adopts the procedure suggested by Demiris et al. (2019), which includes (1) conducting a search, (2) identifying keywords, (3) reviewing the literature, and (4) summarizing and synthesizing findings.

In conducting the search, this study employed several databases, including ScienceDirect, PubMed, MDPI, and Google Scholar. A combination of the following keywords was used: "Kirkpatrick Model", "Training Effectiveness," and "Training Evaluation." The selected articles were journal papers in English published within the last five years (2020-2024) that were relevant to the research topic. Furthermore, in analyzing the data, this paper synthesized relevant findings, concepts, and theories. As Snyder (2019) argued, this process aimed to synthesize the state of knowledge and establish a research agenda for further investigation. According to Jaakkola (2019), theory synthesis "seeks to achieve conceptual integration across multiple theories or literature streams." Thus, this study also sought to elaborate on the limitations of the model and to identify areas for future research beyond the scope of the model.

4.0 Results and Discussion

4.1 The effectiveness of Kirkpatrick's Four-Level Model in Training Evaluation.

The evaluation of the effectiveness of training program can apply to various sectors including industrial, commercial, educational, and governmental organizations or agencies (Morrison et al., 2019). As proposed by Bowman et al. (2020), the effectiveness of evaluation may be based on the consideration of the influence of training interventions, which would involve comparable data and could be achieved through the implementation of a potential approach. The effectiveness of training programs continues to be a topic of contention in the literature, as there is no consensus regarding the most appropriate methodology for evaluating the efficacy of these programs (Bernardino & Curado, 2020). However, Kirkpatrick's Four-Level Model for training evaluation effectiveness has been a cornerstone and most frequently used in assessing the impact of training programs across various industries (Alsalamah & Callinan, 2021).

As outlined by Kirkpatrick (1998), the four-level model – reaction, learning, behavior, and result – represents an evaluation of the effectiveness of the training program. The model offers an assessment of the training's impact, measuring advancements in skills, knowledge, and attitudes, noting performance shifts post-training, and analyzing business outcomes (Urbancová et al., 2021). The effectiveness of the training evaluation is demonstrated in the findings of numerous studies. These findings indicate that the adoption of each level of the model can positively impact the outcome of the training. In this context, the influence of the indicators utilized in the research on the effectiveness of the training at each stage of the training, whether this is the effect on the trainee, the trainer, or the organization, is of interest. Table 1 shows recent studies on the effectiveness of this training evaluation model. The articles under review were published within the last five years.

Table 1: Literature Review of Effectiveness of Kirkpatrick's model of evaluation in different industries

Author/ Year	Research Design	Kirkpatrick's Level	Findings
Bahl et al. (2024)	A quantitative study with PLS-SEM model.	Level 1-4	"Reactions of employees (stage 1) have enhanced knowledge, and skills and has a positive and significant influence on learning (stage 2) and Behavior through job performance (stage 3) has a positive impact on results (stage 4)."
Alordiah (2024)	An explanatory sequential mixed method (QUAN-qual) with survey, interview, and observation of data collection	Level 1-4	"The results showed that high level of participants were satisfied as their reaction of the training, and in learning level, notable advancements in academic writing, publication, and research exposure were observed as a result of the workshop. While in third and fourth level, the participants showed their contribution in article writing, publishing, and research visibility skills and knowledge that indicated the sucess of workshop (result)."
Zhao et al. (2023)	Quantitive metode with data collection using online survey	Level 1-3	"The findings showed that all participants were highly satisfied with the EMI training implementation on the reaction level. On the learning level, participants' scores on awareness of EMI teaching techniques increased significantly with the training proces, and in the behavior level, the participant's confidence as an EMI instructor increased dramatically.
Barile et al. (2023)	An exploratory study based on constructivist grounded theory, with data collection using interview	Level 1-4	"Results show the rise of five core categories, giving rise to an integrated model of Kirkpatrick. Their dynamic interplay led to a new orientation of Kirkpatrick model giving rise to a metalearning ecosystem."
Upadhyay (2023)	Quantitive metode with data collection using survey	Level 1-4	The results showed the high satisfaction of participants in reaction level, and there was a significant increase in participants' knowledge score in learning level, while th emajority of them submit the action plan in behavior level and most of the participant rate the result (level 4) as good.
Quinton et al (2022)	A mixed-method, with data collection using survey.	Level 1- 2	"Higher levels of programme engagement were positively associated with more favourable reactions to the programme. Enjoyment positively predicted learning outcomes, which was mediated by transfer intention."
Bergamo (2022)	A case study with questionnaires administered to the participants.	Level 1-2	"The majority of the participants showed the high satisfaction with the different aspects of training (reaction). Similarly, the study found a high percentage in knowledge performance in terms of practical learning evaluation, but low in theoretical evaluation."
Ramadhan et al. (2022)	An evaluation research methods with a quantitative approach.	Level 1-2	"The result indicated the high positive reaction from the participant's evaluation to implementing the BIM WIKA Allplan Intermediate Training Program, while the learning evaluation found that the training can increase BIM knowledge for most trainees."
Alsalamah and Callinan (2021)	A convergent parallel mixed-methods design, with data collection using survey, semi-strcutured, and literature study.	Level 1-4	"This framework provides a more holistic understanding of the key constructs that ensure the effectiveness of training for female head teachers and head teachers" "The results indicated that the adapted Kirkpatrick evaluation model was very effective in evaluating educational training for head teachers."

Yi et al. (2020)	A quasi-experimental design	Level 1-2	"This study showed a high overall satisfaction and benefit in the reaction level for the IPEP at PUTH, while the higher scores in the learning level did not reach statistical significance."
Sahni (2020)	Quantitive method with data collection using survey	Level 1-2	"The results showed that managerial training using Kirkpatrick Framework is highly effective for both levels; satisfaction and learning."

The utilization of Kirkpatrick's Four-level model from the aforementioned studies is variable, with several studies opting for the first and second levels (Quinton et al., 2022; Bargamo, 2022; Yi et al., 2020; Sahni, 2020; Ramadhan et al. (2022), while another study employed the model up to the third level (Zhao et al., 2023). In contrast, five studies from different industries optimized the model through whole-level evaluations (Bahl et al., 2024; Barile et al., 2023; Alsalamah & Callinan, 2021; Upadhyay, 2023; Alordiah, 2024).

The effectiveness of training evaluations utilizing Kirkpatrick's model, as evidenced by the aforementioned studies, exhibits a dichotomous outcome. The study showed the increased in knowledge, skills and has significantly affect learning to increase job performance and organization performance (Bahl et al., 2024). The model also demonstrated that the effectiveness indicator was used for the outcome in each level, including satisfaction, engagement, enjoyment, knowledge, skills, attitude (Bahl et al., 2024; Quinton et al., 2022; Bargamo, 2022; Yi et al., 2020). The model also idenitifes the key constructs that ensure the effectiveness of the model (Alsalamah and Callinan, 2021), which could form the basis of an integrated framework (Barile et al., 2023). While all studies demonstrated the effectiveness of the selected Kirkpatrick's level employed in the research, two studies (Bargamo, 2022; Yi et al., 2020) revealed that the second level of the model was not entirely effective in evaluating the training. The two studies yielded similar results, with positive findings at the reaction level but insignificant results from the learning level test. For example, the study conducted by Bargamo (2022) demonstrated significant results on the practical test but insignificant results on the theoretical test.

4.2 Limitation of the Model

Despite its numerous advantages, Kirkpatrick's Four-Level Model has faced criticism, with critiques highlighting several theoretical and practical shortcomings (Thalheimer, 2018). For instance, an opposing viewpoint, presented in numerous studies, suggests that the model is inconsistent with research on human learning because it has led to the evaluation of levels 1 and 2 (reaction and learning) by the majority of organizations (Salas et al., 2012; Barile et al., 2023). Although the evaluation of reactions (Level 1) and learning outcomes (Level 2) is of significant importance in gauging the training effectiveness, there is often a discrepancy between the correspondence of these evaluations with tangible behavior modification or organizational ramifications. Participants may offer favorable feedback without manifesting enhanced performance, which can lead to an erroneous perception of effectiveness. Steensma and Groeneveld (2010) argued that this is due to the often unclear connection between training and long-term outcomes. Furthermore, this model is constrained by its assumption of causality among its levels and its significance regarding the rising levels of learning outcomes (Alsalamah & Callinan, 2021). These results are typically influenced by various factors, with training being just one potential factor among many. In essence, the model has deficiencies regarding the importance of contextual factors and fundamental aspects of the program (Lillo-Crespo, et al., 2017). Reio et al. (2017) has proven it by exploring the lack of utilization of the model in the context of job performance and organizational impact of training, as well as the limitations in formative and process evaluations.

Although Kirkpatrick's model is designed to provide a comprehensive evaluation, it is considered resource-intensive due to its complexity, high implementation costs, and time-consuming nature (Jain et al., 2021; Sahni et al., 2020; Housawi et al., 2020). This complexity arises from the different types of data and methodologies involved. Consequently, companies must decide whether the financial investment in the model will lead to enhanced workplace performance and a return on investment from the training (Jain et al., 2021; Urbancová et al., 2021). Reio et al. (2017) noted that while the model is useful, implementing all four levels can be challenging. Cahapay (2021) highlighted the difficulty in directly observing behavioral changes in the workplace, as these changes require concrete evidence beyond participant perceptions typically gathered from surveys. Alordiah (2024) acknowledged this limitation, particularly in accessing data for level four when participants are from different companies or units within a company. As a result, some studies have preferred to focus on measuring training program outcomes at level four, rather than aligning training results with organizational goals (Upadhyay, 2023; Rahman et al., 2019; Alordiah, 2024).

4.3 Beyond the Model

Kirkpatrick's Four-Level Model has served as a fundamental framework for evaluating the effectiveness of training programmes in diverse industries. The limitations of the model, as previously noted, entail significant implications for the effectiveness of evaluation, which may benefit employees involved in the evaluation process. Furthermore, the implications and possibilities for pursuing research on the effectiveness of training evaluation are far-reaching. A clear understanding of the model, together with a relevant instrument for measuring it and addressing its shortcomings, offers the potential to assist in comprehending the more comprehensive and contextual training evaluation measurement and to guide improvement efforts. Training evaluation, as a training process, is a significantly more involved approach to pursue its evaluation effectiveness. Kirkpatrick's Four-Level Model is one of the model that has been effective for some decades but also has difficulties in expensive in its implementation (Johnson & Dick, 2012). Kirkpatrick and Kirkpatrick (2006, p. 18) posit that the most common rationale for evaluation is to "determine the effectiveness of a program and ways in which it can be improved." This paper thus opens the door to an understanding of the conceptual framework of the model and the possibilities of synthesizing the model with other training evaluation approaches for the improvement of future research on comprehensive effective training evaluation.

Kirkpatrick's Model is strongly influended by Reductionist theory, which views knowledge as distinct, identifiable, objective, and impersonal (Allen, 1991). As a result, the model anticipates that training outcomes will enhance the participants' knowledge, skills, and attitudes through the subsequent four levels. The Reductionist theory perceives learning as static and additive, implying that it can be dissected into sequential elements, which these each elements equal as a whole (MacInnis, 1995). This theory is reflected in the model, where Kirkpatrick (1998) argued that the evaluation is conducted in a sequential manner, with the four levels being inextricably linked, yet not occurring simultaneously. In particular, levels 3 (behaviour) and 4 (results) are concerned with long-term measurement (Kirkpatrick & Kirkpatrick, 2006). This introduces a degree of complexity into the evaluation process, given that a multitude of extraneous factors other than the training can influence level 4 outcomes. In light of the complexity of this model, this paper analyzes the main shortcomings of Kirkpatrick's original model, and synthezising them with theories, concepts and evaluations for the development of this model.

As posited by Frye and Hemmer (2012), educational programs, such as training programs, exhibit complexity, involving numerous interactions among participants and the environment. They argue that system theory or complexity theory might be more appropriate for guiding program evaluation.

Consequently, the New World Kirkpatrick Model (NWKM) has been developed as a new version of the original Kirkpatrick's Four Level Model (Kirkpatrick & Kirkpatrick, 2016). The revised model diverges from the original Kirkpatrick framework by prioritizing the sequence of events during the planning phase of training evaluation. This sequence initiates with level 4 and advances to level 1. Following implementation, the levels are assessed in sequential order, starting from 1 and culminating at 4, although this is not necessarily in a linear sequence. Due to this framework, NWKM has been linked to the Complexity Theory (Gandomkar, 2018). In contrast to the linear progression assumed by the traditional Kirkpatrick model, complexity theory acknowledges that training outcomes may not necessarily follow predictable pathways. Instead, minor alterations in the training environment or contextual factors can result in disproportionate effects or tipping points, which in turn influence the effectiveness of training interventions in unforeseen ways. Complexity theory offers researchers a robust and scientifically grounded framework for investigating and comprehending human organizations. It emphasizes acknowledging complexity, patterns, and interconnections instead of merely dwelling on superficial assumptions (Sapir, 2020). In an organizational context, managing employees' careers is complex, so complexity theory provides valuable insights into how to deal with unpredictability and uncertainty in career management systems (Walton, 2016). The complex system within a business organization, as elucidated by Gatrell (2005), is one that is open to feedback from the wider environment within which it exists. The complexity approach may assist evaluators in gaining deeper insights into the contextual factors shaping training effectiveness and in developing more adaptive and holistic evaluation approaches that align with the dynamic nature of training programs within complex organizational systems.

The challenges associated with adopting the traditional four-level model is that it is not context-specific for certain training evaluation settings. The complexity approach has generated the Context evaluation, Input Evaluation, Process evaluation, and Product evaluation (CIPP) model, which deals with contextual aspect of the training evaluation program.

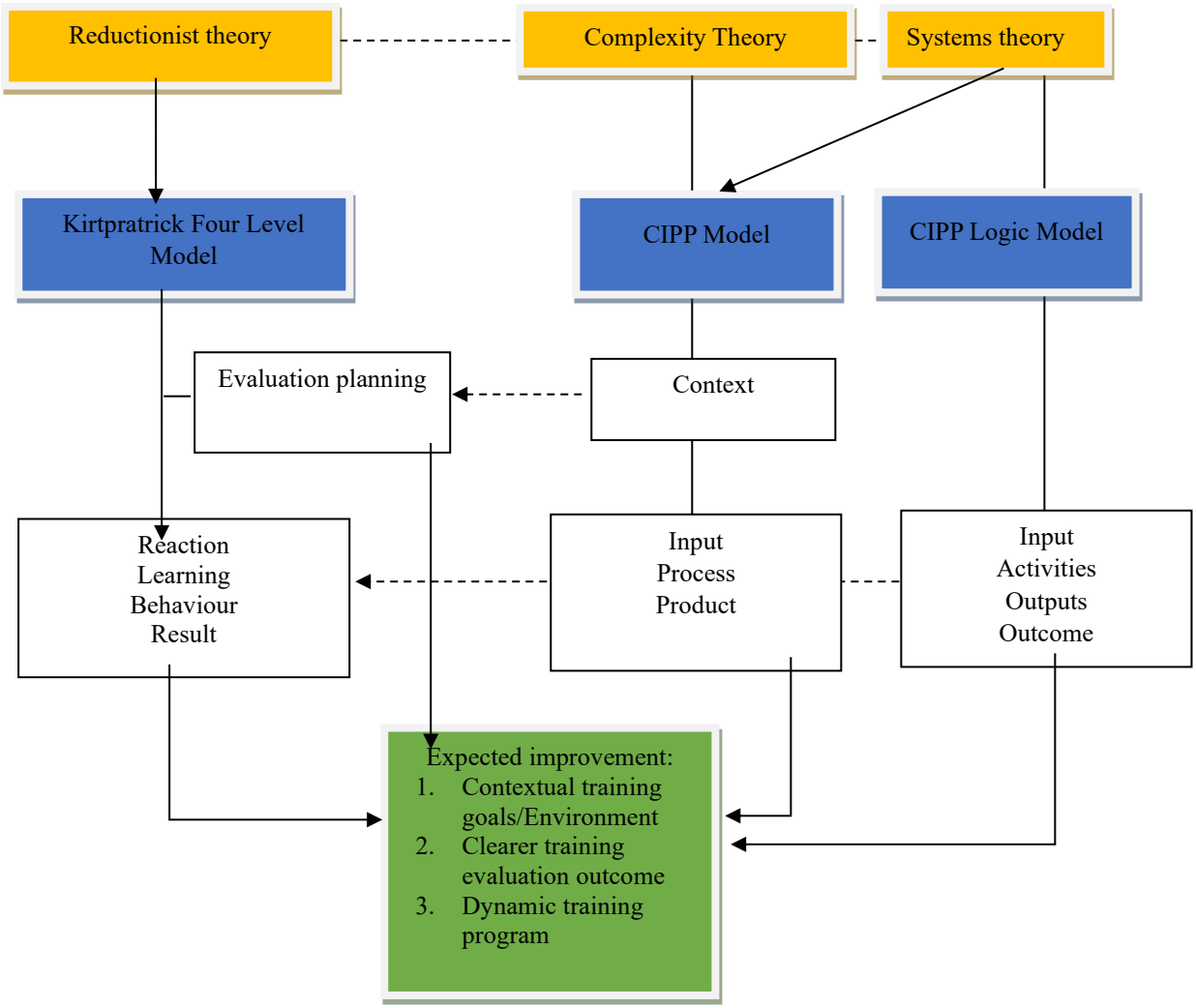

Figure 2: Conceptual Framework

The dynamic interplays between its constituent elements and the external surroundings are acknowledged within the model. As a result, CIPP gives both direct and indirect focus to the program's context, treating contextual evaluation as a distinct facet among four interrelated sets of assessment studies, while also discerning contextual elements across other facets of the model through the utilization of diverse qualitative methodologies (Gandomkar, 2018). Vali and Ataollahi (2022) in their qualitative study over training evaluation argued that CIPP model is a comprehensive, functional model that considers all aspects of the program or system. The context dimension encompasses all crucial elements specific to a particular environment, such as needs, challenges, and opportunities. Meanwhile, the input dimension assesses the strategies, action plans, and resources designated for achieving objectives. The process dimension scrutinizes the actions and documentation required for training completion, while the product dimension identifies both the intended and unintended outcomes of the endeavor (Stufflebeam, 2003). A recent study has reported the effectiveness of the model for training evaluation as it found that employee's understanding level is highest in the training demand, learner's overall understanding level of sub factors of CIPP evaluation model is higher in the satisfaction, and employee's understanding level of outcome of training has bring significant satisfaction to them (Jeon & Lee, 2014).

In terms of contextual aspects, systems theory views organisations as complex systems composed of interconnected parts that influence each other and the system as a whole (Lai & Lin, 2017). Ludwig von Bertalanffy, who proposed the general systems theory, argued that a system is '"a set of elements that stand in relation to each other and to the environment."' (Bertalanffy 1972). Systems theory, which generates the logic model for evaluation, has been widely used for training evaluation effectiveness. The model, which consists of input, activities, outputs and outcome, could be useful for research on Kirkpatrick's four-level model to englight the training outcome difficulties perceived in the model, as the logic model is explisictly focus on outcome. For example, Damasceno et al. (2012) investigated the effectiveness of the Logic Model on organisational performance. Another study showed the importance of the model for performance measurement (Besharov & Call, 2016).

5.0 Conclusion

This paper provides an in-depth exploration of training evaluation, with a particular focus on defining the concept of Kirkpatrick's Four-Level Model and outlining the effectiveness in training evaluation from adopting the model. In conclusion, Kirkpatrick's Four-Level Model is still relevant for use in examining the effectiveness of training evaluations. The relevance of this model can be seen from the accuracy of the model indicators used, despite the inability of this model to be implemented contextually in all organizations. This is evident in the utilization of various levels and the contextual discrepancy in employing levels 3 and 4 of the model. Thus, the New World Kirkpatrick Model (NWKM) serves as a preferred framework for future research to rectify the limitations of the original model. In addition, theoretical foundations related to Training Evaluation, such as Complexity Theory and System Theory and their derived evaluation models are suggested to be incorporated into research using the Kirkpatrick Model to underpin the key constructs of the model and optimize the overall outcome of evaluation results.

References

Alordiah, C. O. (2024). Evaluation of a research training workshop for academic staff in tertiary institutions: A Kirkpatrick model approach. *Journal of Applied Learning and Teaching, 7*(1), 303-313.

Alsalamah, A., & Callinan, C. (2021). Adaptation of Kirkpatrick's four-level model of training criteria to evaluate training programmes for head teachers. *Education Sciences, 11*(3), 116.

Al Qudah, N. F., Yang, Y., & Anjum, M. A. (2018). Transformational training programs and quality orientation of employees: Does employees' loyalty matter?. *Sustainability, 10*(2), 465.

Asari, M. S., & Leman, A. M. (2015). Safety training evaluation: Approaches and practices. *Journal of Occupational Safety and Health, 12*(2), 23-30

Asghar, M. Z., Afzaal, M. N., Iqbal, J., Waqar, Y., & Seitamaa-Hakkarainen, P. (2022). Evaluation of in-service vocational teacher training program: A blend of face-to-face, online and offline learning approaches. *Sustainability, 14*(21), 13906.

Bahl, K., Kiran, R., & Sharma, A. (2024). Evaluating the effectiveness of training of managerial and non-managerial bank employees using Kirkpatrick's model for evaluation of training. *Humanities and Social Sciences Communications, 11*(1), 1-14.

Barile, S., Ciasullo, M. V., Testa, M., & La Sala, A. (2023). An integrated learning framework of corporate training system: a grounded theory approach. *The TQM Journal, 35*(5), 1106-1134.

Bates, R. (2004). A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Evaluation and program planning, 27*(3), 341-347.

Bergamo, P. A. D. S., Streng, E. S., de Carvalho, M. A., Rosenkranz, J., & Ghorbani, Y. (2022). Simulation-based training and learning: A review on technology-enhanced education for the minerals industry. *Minerals Engineering, 175*, 107272.

Bernardino, G., & Curado, C. (2020). Training evaluation: a configurational analysis of success and failure of trainers and trainees. *European Journal of Training and Development, 44*(4/5), 531-546.

Besharov, D. J., & Call, D. M. (2016). Using Logic Models to Strengthen Performance Measurement. In D. J. Besharov, K. J. Baehler, & J. A. Klerman (Eds.), *Improving Public Services: International Experiences in Using Evaluation Tools to Measure Program Performance* (pp. 303–324). Oxford Academic.

Bęś, P., & Strzałkowski, P. (2024). Analysis of the Effectiveness of Safety Training Methods. Sustainability, 16(7), 2732.

Bowman, S. R., Palomaki, J., Soares, L. B., & Pitler, E. (2020). Collecting entailment data for pretraining: New protocols and negative results. *In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 8203-8214).

Brinia, V., & Efstathiou, M. (2012). Evaluation of factors affecting training transfer on safety in the workplace: a case study in a big factory in Greece. *Industrial and Commercial Training, 44*(4), 223-231.

Brown, K. G. (2007). Training Evaluation. In S. G. Rogelberg (Ed.), *Encyclopaedia of Industrial and Organisational Psychology* (pp. 820–823). Thousand Oaks, California: SAGE Publications, Inc.

Cahapay, M. (2021). Kirkpatrick model: Its limitations as used in higher education evaluation. *International Journal of Assessment Tools in Education*, *8*(1), 135-144.

Chao, J., Luo, P., Yeh, Y., & Kao, H. (2018). The study of learning effects of mobile devices courses for Taiwanese senior citizens using the Kirkpatrick model. *The International Journal of Electrical Engineering & Education*, 0020720918816003.

Damasceno, S. M. S., Abbad, G., & Meneses, P. P. M. (2012). Logic models and organizational training evaluations. *Paidéia (Ribeirão Preto), 22*, 217-228.

Demerouti, E., & Peeters, M. C. (2018). Transmission of reduction-oriented crafting among colleagues: A diary study on the moderating role of working conditions. *Journal of Occupational and Organizational Psychology, 91*(2), 209-234.

Demiris, G., Washington, K. T., & Oliver, D. P. (2019). *Behavioral Intervention Research in Hospice and Palliative Care Building an Evidence Base*. Elsevier

El Nsouli, D., Nelson, D., Nsouli, L., Curtis, F., Ahmed, S. I., McGonagle, I., ... & Ahmadi, K. (2023). The application of Kirkpatrick's evaluation model in the assessment of interprofessional simulation activities involving pharmacy students: a systematic review. *American Journal of Pharmaceutical Education, 8*, 100003

Ford, J. K. (2007). Transfer of Training. In S. G. Rogelberg (Ed.), *Encyclopaedia of Industrial and Organisational Psychology* (pp. 832-835). Thousand Oaks, California: SAGE Publications, Inc.

Frye, A. W., & Hemmer, P. A. (2012). Program evaluation models and related theories: AMEE guide no. 67. *Medical teacher, 34*(5), e288-e299.

Gebrehiwot, G. D., & Elantheraiyan, P. (2023). A study on the effect of training on employee performance in the case of Mekelle City, Tigray, Ethiopia. *Social Sciences & Humanities Open, 8*(1), 100567.

Giday, D. G., & Elantheraiyan, P. (2023). A study on the effect of training on employee performance in the case of Mekelle City, Tigray, Ethiopia. *Social Sciences & Humanities Open, 8*(1), 100567.

Gil, A. J., Rodriguez-Cavides, L., & Romero-Daza, D. (2023). Analysis of training effectiveness from the perspective of managers and employees in the Colombian hospitality industry. *Industrial and Commercial Training, 55*(3), 346-354.

Goh, W. W., Wong, S. Y., & Ayub, E. (2018). The effectiveness of MOOC among learners based on Kirkpatrick's model. *In Redesigning Learning for Greater Social Impact: Taylor's 9th Teaching and Learning Conference 2016 Proceedings* (pp. 313-323). Springer Singapore.

Gregory, A. T., & Denniss, A. R. (2018). An introduction to writing narrative and systematic reviews—Tasks, tips and traps for aspiring authors. *Heart, Lung and Circulation, 27*(7), 893-898.

Housawi, A., Al Amoudi, A., Alsaywid, B., Lytras, M., bin Moreba, Y. H., Abuznadah, W., & Alhaidar, S. A. (2020). Evaluation of key performance indicators (KPIs) for sustainable postgraduate medical training: an opportunity for implementing an innovative approach to advance the quality of training programs at the Saudi Commission for Health Specialties (SCFHS). S*ustainability, 12*(19), 8030.

Jain, G., Sharma, N., & Shrivastava, A. (2021). Enhancing training effectiveness for organizations through blockchain-enabled training effectiveness measurement (BETEM). *Journal of Organizational Change Management, 34*(2), 439-461.

Jeon, J., & Lee, J. (2014, 03-06 December). *A Study on Satisfaction of Employee Training Program CIPP Evaluation Model Focusing on A Corporation.* 2014 International Conference on Interactive Collaborative Learning (ICL), Dubai, UAE.

Karim, A. A., Abduh, A., Manda, D., & Yunus, M. (2018). The Effectivity of Authentic Assessment Based Character Education Evaluation Model. *TEM journal, 7*(3), 495-500.

Khalid, M., Ashraf, M., & Rehman, C. A. (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.

Khandker, S. R., Koolwal, G. B., & Samad, H. A. (2010). *Handbook on Impact Evaluation: Quantitative Methods and Practices.* The International Bank for Reconstruction and Development / The World Bank.

Kirkpatrick, D.L. (1994). *Evaluating Training Programs: The Four Levels.* Berrett-Koehler.

Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels*. 2nd Ed. Berrett-Koehler.

Kirkpatrick, D. L., & Kirkpatrick, J. D (2006). *Evaluating training programs: The four levels*. 3rd Ed. Berrett-Koehler

MacInnis, C. (1995). Holistic and reductionist approaches in special education: Conflicts and common ground. *McGill Journal of Education/Revue des sciences de l'éducation de McGill, 30*(001), 7-20.

Moldovan, L. (2016). Training outcome evaluation model. *Procedia Technology, 22*, 1184-1190.

Noe, R. A. (2017). Employee Training and Development. 5th Ed. McGraw-Hill Education.

Pedrini, G., & Cappiello, G. (2022). The impact of training on labour productivity in the European utilities sector: An empirical analysis. *Utilities Policy, 74,* 101317.

Phillips, J. J., & Phillips, P. P. (2016). *Handbook of training evaluation and measurement methods*. Routledge.

Priyono, M. I., & Perkasa, D. H. (2024). Determinan Faktor Pengembangan Karir Karyawan: Pengalaman Kerja, Tingkat Pendidikan Dan Karakteristik Individu. Jurnal Bina Bangsa Ekonomika, 17(2), Irfan-Priyono.

Quinton, M. L., Tidmarsh, G., Parry, B. J., & Cumming, J. (2022). A Kirkpatrick model process evaluation of reactions and learning from my strengths training for life™. *International Journal of Environmental Research and Public Health, 19*(18), 11320.

Ramadhan, M. A., Sumarsono, R. A., Achmad, A. A., & Cisse, A. (2022). Implementation of Kirkpatrick evaluation model in Building Information Modeling (BIM) training program. *Jurnal Pendidikan Teknologi dan Kejuruan, 28*(2), 130-138.

Reio, T. G., Rocco, T. S., Smith, D. H., & Chang, E. (2017). A Critique of Kirkpatrick's Evaluation Model. *New Horizons in Adult Education and Human Resource Development, 29*(2), 35–53.

Sapir, J. (2019). *Thriving at the edge of chaos: Managing projects as complex adaptive systems*. Productivity Press.

Sahni, J. (2020). Managerial training effectiveness: An assessment through Kirkpatrick framework. *TEM Journal, 9*(3), 1227-1233.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of business research, 104,* 333-339.

Steensma, H., & Groeneveld, K. (2010). Evaluating a training using the "four levels model*". Journal of Workplace Learning, 22*(5), 319-331.

Upadhyay, K., Goel, S., & John, P. (2023). Developing a capacity building training model for public health managers of low and middle income countries. *PloS one, 18*(4), e0272793

Urbancová, H., Vrabcová, P., Hudáková, M., & Petrů, G. J. (2021). Effective training evaluation: The role of factors influencing the evaluation of effectiveness of employee training and development. *Sustainability, 13*(5), 2721.

Vali, L., Ataollahi, F., Amiresmaili, M., Nakhaee, N., & Okhovati, M. (2022). The requirements of developing programs for the management of non-communicable diseases in Iran based on the CIPP model: A qualitative study. *Journal of Health Research, 36*(6), 1068-1077.

Walton, M. (2016). Setting the context for using complexity theory in evaluation: boundaries, governance and utilisation. *Evidence & Policy, 12*(1), 73-89.

Yi, Z. M., Zhou, L. Y., Yang, L., Yang, L., Liu, W., Zhao, R. S., & Zhai, S. D. (2020). Effect of the international pharmacy education programs: A pilot evaluation based on Kirkpatrick's model. *Medicine, 99*(27), e20945.

Zhang, L., Guo, X., Lei, Z., & Lim, M. K. (2019). Social network analysis of sustainable human resource management from the employee training's perspective. *Sustainability, 11*(2), 380.

Zhao, W., Liu, Z., Wang, T., Yin, X., Sun, Y., Zhang, X., & Yang, H. (2023). Assessment of a training project of English as a media of instruction(EMI) using Kirkpatrick model. *BMC medical education, 23*(1), 271.

Pejabat Perpustakaan
Library Office

Universiti Teknologi MARA
Cawangan Perak
Kampus Seri Iskandar
32610 Bandar Baru Seri Iskandar,
Perak Darul Ridzuan, MALAYSIA
Tel: (+605) 374 2093/2453 Faks: (+605) 374 2299

UNIVERSITI TEKNOLOGI MARA

Surat kami : 700-KPK (PRP.UP.1/20/1)
Tarikh : 20 Januari 2023

Prof. Madya Dr. Nur Hisham Ibrahim
Rektor
Universiti Teknologi MARA
Cawangan Perak

Tuan,

PERMOHONAN KELULUSAN MEMUAT NAIK PENERBITAN UiTM CAWANGAN PERAK MELALUI REPOSITORI INSTITUSI UiTM (IR)

Perkara di atas adalah dirujuk.

2. Adalah dimaklumkan bahawa pihak kami ingin memohon kelulusan tuan untuk mengimbas (*digitize*) dan memuat naik semua jenis penerbitan di bawah UiTM Cawangan Perak melalui Repositori Institusi UiTM, PTAR.

3. Tujuan permohonan ini adalah bagi membolehkan akses yang lebih meluas oleh pengguna perpustakaan terhadap semua maklumat yang terkandung di dalam penerbitan melalui laman Web PTAR UiTM Cawangan Perak.

Kelulusan daripada pihak tuan dalam perkara ini amat dihargai.

Sekian, terima kasih.

"BERKHIDMAT UNTUK NEGARA"

Setuju.

27.1.2023

PROF. MADYA DR. NUR HISHAM IBRAHIM
REKTOR
UNIVERSITI TEKNOLOGI MARA
CAWANGAN PERAK
KAMPUS SERI ISKANDAR

Saya yang menjalankan amanah,

SITI BASRIYAH SHAIK BAHARUDIN
Timbalan Ketua Pustakawan

nar