JTH

https://jthkkss.com/
e-ISSN 2805-4431
DOI: https://doi.org/10.53797/jthkkss.v5i1.4.2024

Evaluation Models in Curriculum and Educational Program - A Document Analysis Research

Liu, Shu[1,2*] & Zu, Yanbin[1,2]

[1]Zhumadian Vocational and Technical College, Zhumadian, 463000, CHINA
[2]Faculty of Technical and Vocational Education, Universiti Tun Hussein Onn Malaysia, MALAYSIA

*Corresponding author email: zzyliushu@gmail.com

Received 03 February 20240; Accepted 01 July 2024; Available online 06 July 2024

Abstract: An educational or instructional program's efficacy may be impacted by the curriculum, which is a significant factor. The issue of evaluating curriculum and instruction has been of concern to researchers at every time and every place. As research progresses, various evaluation models have emerged and are being revised and refined. The purpose of this research is to review some of the best-known models for evaluating curriculum instruction. During this process, researchers used document analysis methods to select relevant research documents based on predetermined criteria. These documents included articles and books. These models are then discussed. In addition, the advantages and disadvantages of some models are shown. It is concluded that the appropriate models or the necessary combinations need to be selected according to the actual requirements of the project. Finally, the researchers explore an evaluation model based on previous research to evaluate student curriculum learning outcomes.

Keywords: Evaluation models; Curriculum evaluation; Educational program evaluation; Document analysis

1. Introduction

The core task of TVET development in the IR4.0 era is to improve the quality of talent development (Liu, Wang & Wang, 2016), and the TVET curriculum significantly affects whether educational programs are successful or unsuccessful (Drakos, 2005). The curriculum incorporates what should be delivered in an educational environment, along with a predefined set of instructional methods, materials, and evaluation standards (Medgyes & Nikolov, 2000). The curriculum gives participants a clear understanding of what needs to be accomplished throughout project implementation and how things actually turned out at the end of the project (Drakos, 2005). According to Cai (2003), effective curriculum instruction includes curriculum design and curriculum development, classroom performance, and assessment of curriculum instruction. The strengthening and improvement of curriculum development and assessment is the key to promoting teaching and learning infrastructure and improving teaching quality and faculty standards. As a result, curriculum assessment has become a crucial component of raising the standard of both teaching and learning as well as the standard of human resources. However, there is no one single criterion is sufficient to assess effective teaching and learning (Marsh, 1987).

Curriculum assessment is an important component of most assessment programs, especially educational programs, and is an essential phase of educational curriculum and program development (Nouraey et al., 2020). However, the appropriateness of evaluation models depends on the context, and evaluators are faced with the task of adapting the models (McNamara, 2000). This is where the use of certain evaluation models, rather than the models themselves, presents serious limitations (Cahapay, 2021).

The primary goal of program evaluation is to determine if the program is effective or whether changes need to be made (Rossi, Lipsey & Henry, 2018). Some academics have a propensity to systematically classify the potential goals of program evaluation. For example, Kirkpatrick (1994) proposed a four-level training evaluation model as a standard training program evaluation tool, and this model is widely used to evaluate the effectiveness of training programs. SCM

*Corresponding author: zzyliushu@gmail.com

evaluates the effectiveness of training by intentionally looking for the best results produced by the training (Brinkerhoff, 2005). The Technology Acceptance Model (TAM), proposed by Fred Davis, is widely used to examine the factors affecting user acceptance of technology. In the context of education, conducting program evaluation implies both professional responsibility and program development (Nouraey et al., 2020).

To date, a large number of researchers have tended to elucidate research on the development of assessment models (Rossi, Lipsey & Henry, 2019; Brinkerhoff, 2003; Kirkpatrick, 1994; Philips, 2012; Stufflebeam, 2003). There are already more than 50 models for evaluating curriculum (Kavgaoglu & Alci, 2016). Several of these models are prominent in the field and have been continuously explored by other academics, while others have only been examined occasionally (Nouraey et al., 2020). Therefore, the purpose of this research is to review some prominent program and course instructional assessment models and examine their concepts, theories, and characteristics. This research was conducted to help researchers quickly and clearly enter this field of study and to provide guidance in selecting the correct and appropriate models for future evaluation efforts, including instructional programs, courses, and teaching. Generally, the research objectives of this study include:

1. Identifying five well-known models for assessing curriculum and educational program.
2. Identify the concepts and characteristics of these five models.
3. Attempting to develop a curriculum instructional assessment model based on these models.

2. Methodology

According to Bowen (2009), document analysis is a methodical approach to reading and assessing literature, in which researchers interpret the literature to give voice and meaning around the topic being assessed. Literature analysis can include both quantitative and qualitative components, with emphasis here on the qualitative approach, which is common as a research method in many social science disciplines. O'Leary (2014) classifies documents as having three main types, namely Public Records, Personal Documents, and Physical Evidence. This research used a document analysis method (Bowen, 2009) in which Google Scholar, a database of literature, was used as a data source and the keywords "Educational Evaluation Model" were used to search for papers that were included based on various criteria, incorporating pertinent information about the search terms, the research journal's impact factor, and the quantity of citations. Number of citations. After the search, the researchers read the titles, abstracts, and keywords of the articles to initially obtain articles related to the topic of the study. An in-depth reading of the identified documents was then performed.

3. Findings & Discussion

3.1 Rossi's Five Domain Evaluation model

In the late 1970s, Rossi et al. developed an evaluation model called the Five Domain Evaluation Model (Rossi et al., 2019). In this model, each assessment should be customized based on local needs, resources, and program type. This includes customizing the assessment questions, methods and procedures, and the nature of the assessor-stakeholder relationship.

The Five domain evaluation model highlights five main evaluation domains.

1. Program needs assessment
2. Program design and theory
3. Program implementation and process assessment
4. Program outcome or impact assessment
5. Program cost and efficiency assessment

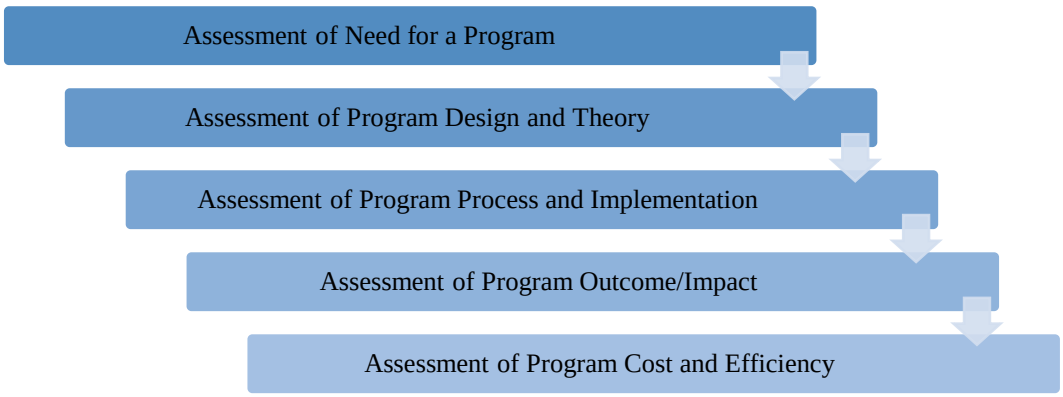

Figure 1: Rossi's Five Domains of Evaluation

In this evaluation model, each domain builds on the previous one. Ideally, programs will not be designed and implemented unless the need for the program is identified. If the program is not implemented, it will also not be possible

to evaluate program results. It is not appropriate to conduct an efficiency evaluation without determining whether the program is achieving the desired results.

3.2 Brinkerhoff's Success Case Method

The Success Case Method (SCM) (Brinkerhoff, 2003) entails finding and carefully assessing the greatest and least successful cases in a program. This approach was developed by Brinkerhoff to evaluate the impact of organizational interventions, such as training and coaching, although the use of SCM is not limited to this context. It is an effective technique for compiling impact accounts and for gaining a deeper comprehension of the elements that either strengthen or weaken impact.

The greatest and least successful program participants are specifically highlighted in the success case model. Instead of focusing on average performance, the aim is to uncover and analyze extreme cases to answer the question, "How well does the program operate when it runs? What is effective and ineffective?"

SCM research can be used to answer any or all of the following questions:

"What really happened?

What results, if any, did the program help produce?

What is the value of those results?

How can the initiative be improved?"(Brinkerhoff, 2003, p. 6-7).

Direct results of conducting success case method studies include documented impact stories that can be disseminated to stakeholders and a better understanding of the factors that enhance or impede business impact.

In the SCM approach, there are five key steps (Brinkerhoff, 2003, p. 29)

1. Sharpen the focus and plan a successful case study
2. Create an "impact model" that defines what success should look like
3. Design and implement a survey to find the best and worst cases
4. Interview and document success stories
5. Communicate findings, conclusions and recommendations

3.3 The Four-Level Evaluation Model

Kirkpatrick's Four-level model of evaluation for learning, which was first suggested in the 1950s, is a thorough approach of curriculum and program evaluation. (Kirkpatrick & Kirkpatrick, 2006). The model has undergone a number of adjustments since then, but the fundamental ideas (the four basic evaluation levels) are still valid today (Hamemoradi, 2014).

In Kirkpatrick's opinion, the information or data collected for an assessment can generally be divided into four different categories or levels. That is, an evaluation can typically consist of four different levels of platforms.

1. "Reactions" platform. Reaction is the evaluative feedback of the participants on the training program and its results based on their own impressions and feelings. The "reaction" evaluation is conducted in order to understand the participants' views and opinions on the instructors, teaching materials, teaching methods and approaches, as well as the organization and management of teaching. In general, only a training program and implementation process that is generally satisfactory to the participants is likely to provide a fuller picture of the actual effectiveness of the training.

2. "Learning" platform. This level is for evaluating the trainees' educational activities. Its primary objective is to assess the trainees' cognitive and technical proficiency, as well as their level of post-training mastery of the related concepts, information, abilities, and methods. This level of evaluation methods, including oral examinations, written tests, field operations and job simulations, etc.

3. "Behavior" platform. By behavior, we mean the work behavior of employees. Behavior change evaluation is to determine the degree of change in the actual work behavior of the trainees after they return to their jobs after training. In other words, it is to measure the extent to which the knowledge and skills learned by the trainees have been translated into improvement and enhancement of actual work behavior. Behavior change can be evaluated in the areas of work attitude, work methods, productivity, attendance, scrap rate, and accident rate, etc. The evaluation methods can be self-assessment, colleague evaluation, and supervisor evaluation, etc.

4. "Results" platform. This level is for evaluating the training for enhancing organizational performance. According to Kirkpatrick (2006), following training, production cost savings, output changes, and quality improvements can all be tracked through analysis and evaluation of improvement in organizational performance.

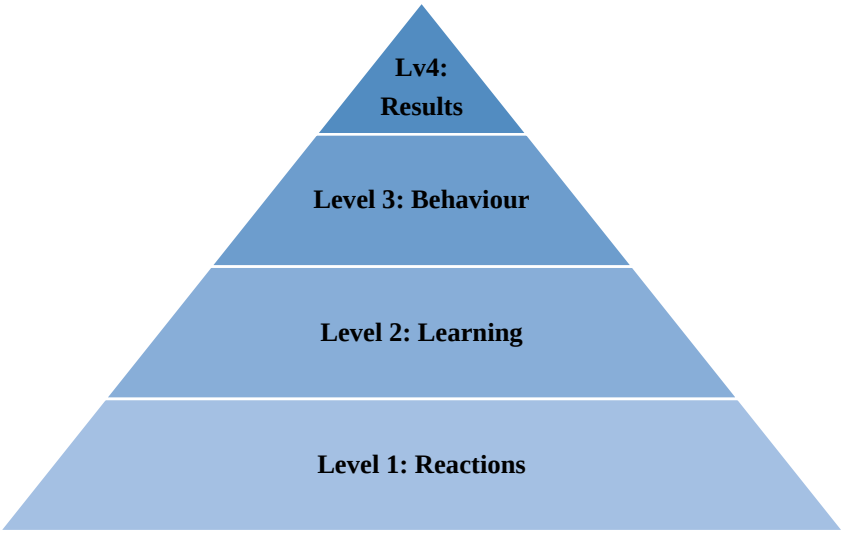

Figure 2: The Four-Level Evaluation Model

From this, Kirkpatrick's "Four Levels of Evaluation Model" aims to answer four important questions:

1. Whether and in what ways are the participants satisfied with the training program?
2. What did the trainees master as a result of the training program?
3. Did the training alter the participants' behavior?
4. Whether the change in behavior has a positive impact on organizational performance?

The actual situation shows that these four questions are the most instinctive and expected answers for any person or organization related to training activities, and they are also the fundamental questions that can most directly reflect the value and meaning of training activities and training evaluation itself. Therefore, Kirkpatrick's " Four Level Evaluation Model" is often referred to as the most classic and well-known training evaluation model and is widely used in training evaluation practice in many countries.

3.4 Philips' Model of Learning Evaluation

Although Kirkpatrick's " Four Levels Evaluation Model" is known for its classic meaning, flaws were quickly identified, and attempts were made to revise this basic evaluation model. The Philips' evaluation model (2012) (Table 1), additionally referred to as the learning evaluation model, is thought to be a supplementary model to Kirkpatrick's four levels of evaluation by including a fifth level: Return on Investment (ROI) to the model. In summary, Philips' model is concerned with collecting data, separating out the effects of training from other factors, and considering additional benefits. The core task is to compare the monetary profit generated by the training program with its cost.

Table 1: Philips' Model of Learning Evaluation (Nouraey et al., 2020)

Levels of Evaluation	Measurement Focus	Time Frame
Level 1: Reaction	Participant reaction to a learning program.	Conclusion of program
Level 2: Learning	Degree to which participants acquired new knowledge, skills or attitudes.	Conclusion of program or within 6 to 8 weeks after
Level 3: Behavior	Degree to which participants applied back-on-the-job what was learnt.	2 to 12 months
Level 4: Results	Degree to which targeted business outcomes were achieved.	9 to 18 months
Level 5: ROI	Degree to which monetary program bene fits exceed program costs.	9 to 18 months

Philips (2012) notes that many HRD organizations place great emphasis on training evaluation, but they are more concerned with general satisfaction with training and rarely assess "ROI". According to Philips (2012), the "ROI" evaluation may indeed be a very difficult, complex and well-planned exercise. However, it is only when this level of evaluation is completed that the entire training evaluation process can be declared truly complete.

Phillips' revision of Kirkpatrick's model, in addition to adding the crucial "ROI" platform, also provides tips on what to look for at each level of evaluation. He points out that: The first level of checking participant satisfaction must acknowledge that positive feedback does not always mean that participants have acquired new knowledge and skills; the

second level of learning evaluation must be aware that the positive results about the platform, not always guaranteed to be applied in the workplace; the third level is to assess the application of training content in the workplace, it is important to note that the training program does not necessarily. It is crucial to keep in mind that training initiatives may not always benefit the organization in a positive way; as a result, they must be carefully and impartially evaluated; the fourth level is to evaluate the impact of the knowledge applied by the participants on the performance of the organization, and it is important to recognize that training programs may produce significant performance, but its cost investment will have a significant impact on performance, so the last level of "ROI " evaluation.

Through Phillips' tips for each level of evaluation, one can discover that the entire training evaluation process is actually a series of connected processes: the first and most fundamental link is the evaluation of satisfaction; the second is the evaluation of learning activities, which lays the groundwork for testing the application of newly acquired knowledge and skills; the third is the evaluation of knowledge and skill application, which is a requirement for evaluating organizational performance following training; and the fourth is the evaluation of organizational performance changes, which is the only way to determine the return on investment.

However, fundamentally, Phillips' evaluation model is a more traditional systemic training model of evaluation. The basic components of this training model are determining training needs → developing training plans → implementing training activities → evaluating training results. The step of "evaluating training results", which is the last part of the whole training process, is basically irrelevant to the other steps before it, except that it may have some influence on the subsequent training programs. The evaluation model is basically designed to accomplish the task of the last phase in the training model. As a result, these evaluation models are largely of the "summative evaluation" variety. This leads to one of its shortcomings, that is, all the efforts made for training evaluation, are almost useless for the training activity itself - neither to talk about the implementation of the necessary monitoring of its entire process, nor to make it possible to continuously obtain the necessary adjustments and improvements. Thus, the "CIPP" training evaluation model, which advocates evaluation activities throughout the training process and has obvious "process" and "formative" characteristics, has been developed.

3.5 The CIPP Model

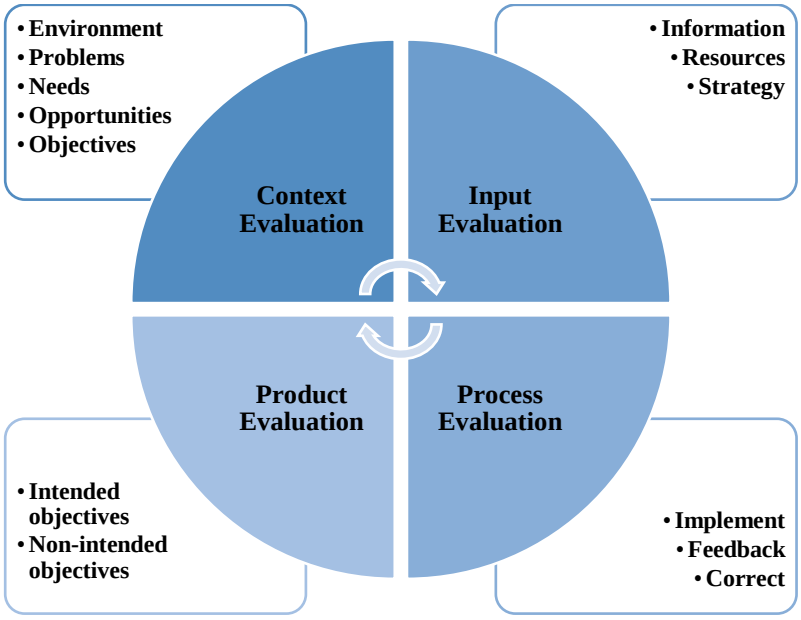

Figure 3: The CIPP Model

The CIPP model (Stufflebeam, 2003) is a comprehensive, detailed evaluation model that focuses on four major areas of the project. Stufflebeam defines the components of the model as follows:

1. Context evaluation: Understanding the pertinent environment, diagnosing unique issues, analyzing training needs, identifying training needs, identifying training opportunities, setting training objectives, etc. The primary tasks among them are identifying training needs and establishing training objectives.
2. Input evaluation: Assemble data on available training resources; assess those resources; decide how to best utilize those resources to meet training goals; and ascertain whether the overall approach to project planning and design necessitates the use of outside resources.
3. Process evaluation: Insight into the potential reasons why training initiatives fail, along with programs that aim to address those reasons; analyze the negative factors that lead to training implementation failure and suggest solutions; assess and explain what actually occurs during the implementation of the training and the

circumstance; evaluate and assess the distance between them and the goals; insist on giving details regarding decisions made during the implementation of training, both new and existing decisions, etc. It is true that process assessment requires a substantial amount of pertinent data to be supported, much like other evaluation phases. Formal and informal techniques, such as the use of rating scales, feedback forms, and record analysis, can be used to gather this information and data. Those in charge of putting training programs into action should get feedback from process evaluation in order to make prompt and continual corrections or improvements to the way training programs are put into practice.

4. Product evaluation: The measuring and interpretation of the goals attained through training activities, including the measurement and interpretation of both the goals attained that were intended and the goals attained that were unintended.

The CIPP evaluation model has very distinctive features, especially its full, process and feedback.

Full: The evaluation activity is actually connected to every step of the training activity, or rather, it is truly integrated into every step of the training process. For example, the phases "identification of training needs" and "determination of training objectives" are related to "background evaluation," "input evaluation" is related to the steps "deciding training strategy" and "designing and planning training," and "process evaluation" is related to the phase "implement training."

Process: It is suggested that the training program's implementation process be watched carefully so that any potential failure factors, unfavorable circumstances, and the distance between implementation and training objectives can be identified. The training program should also be guided through the implementation process to make timely and appropriate strategic and tactical adjustments or suggestions for ways to improve.

Feedback: According to the CIPP approach, "outcome evaluation" can be carried out both before and after training. In other words, the CIPP model wants to conduct "outcome evaluation" both before and after the training so that its feedback can be more beneficial to the subsequent training programs. It also wants to conduct "outcome evaluation" while the training is taking place so that its feedback can be more beneficial to the ongoing training activities. As demonstrated by experience, "outcome evaluation" in training implementation will once more provide a more useful basis and motivation for improving and facilitating the training process on the one hand, and will support fully utilizing the learning potential of the participants and boosting their motivation on the other.

4. Conclusion & Suggestion

Scholars and researchers in the area have proposed a variety of program and course evaluation models. Some of these models have been continuously tested, while some are less commonly mentioned. The ultimate goal of all these evaluation models is to figure out whether a program or course is meeting its goals. Each evaluation model has strengths and weaknesses. For example, Kirkpatrick' Four Levels and Phillips' Five Levels evaluation model, while lacking consideration of all aspects of the training process, has a detailed and operational design for post-training evaluation; The CIPP model illustrates the importance of the entire process, including the process and feedback from the evaluation, but it also leaves room for improvement in terms of "outcome evaluation" and other areas. In this sense, their relationship is a "complementary" one. When one needs to apply them in practice, it is necessary to select them appropriately or combine them as necessary according to the actual requirements of the organization or the training program. These suggestions might result in assessments that are more productive and efficient. Therefore, the researchers suggested a comprehensive evaluation of course learning effectiveness of students in Zhumadian Vocational and Technical College based on the CIPP model and the Kirkpatrick' Four-Level Evaluation Model. The researchers divided the course learning effectiveness evaluation into three stages, corresponding to before, during, and after the class. The background and input evaluations in the CIPP model were incorporated into the pre-course stage. Considering that course learning is a cyclical process, the researchers incorporated the Kirkpatrick' model's response, learning, and behavioral evaluation indicators into the process evaluation: the during class phase, in order to promote learning quality during the learning process. Product evaluation was implemented at the end of the course to evaluate the effectiveness of student learning.

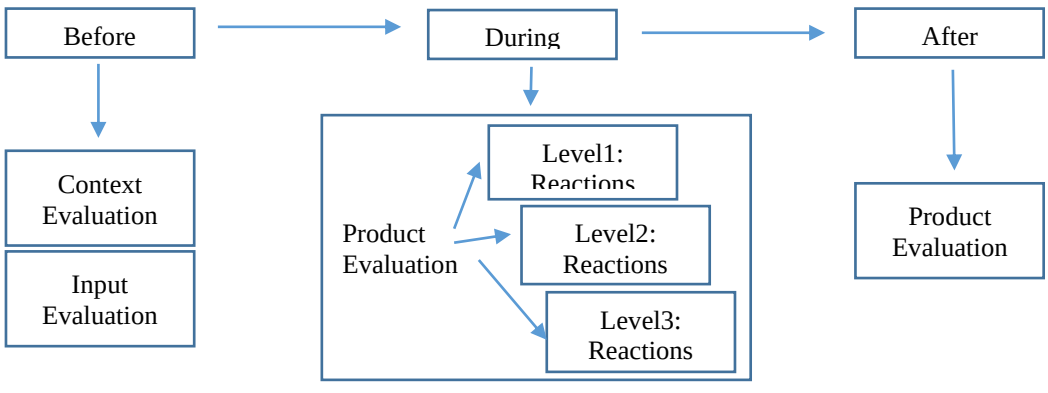

Figure 4: The Modified CIPP & Kirkpatrick' Four-Level Evaluation Model

References

Bowen, G. A. (2009). Document Analysis as a Qualitative Research Method. Qualitative Research Journal, 9(2), 27-40.

Brinkerhoff, R. O. (2003). The Success Case Method: Find Out Quickly What's Working and What's Not. Berrett-Koehler Publishers.

Brinkerhoff, R. O. (2005). The Success Case Method: A Strategic Evaluation Approach to Increasing the Value and Effect of Training. Advances in Developing Human Resources, 7(1), 86-101.

Cahapay, M. (2021). Kirkpatrick Model: Its Limitations as Used in Higher Education Evaluation. International Journal of Assessment Tools in Education, 8(1), 135-144.

Cai, Q. H. (2003). Focus on Curriculum Evaluation and Promote Curriculum Construction. Journal of Shaoyang University (Social Science Edition), 2(03), 135-137.

Drakos, J. (2005). Ideas for developing a personal EFL teaching curriculum. The Internet TESL Journal, XI (3), 1-3.

Hamemoradi, M., Khorasani, A., & FathiAjargah, K. (2014). Evaluation of the Effectiveness of the on-the-job Training Courses in NIGC Based Kirkpatrick, Phillips and CIPP Models. Strategic Studies in Petroleum and energy Industry, 5(19), 103-130.

Kavgaoglu, D., & Alci, B. (2016). Application of Context Input Process and Product Model in Curriculum Evaluation: Case Study of a Call Centre. Educational Research and Reviews, 11(17), 1659-1669.

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

Liu, X. Y., Wang, Z. B., & Wang, Z. Y. (2016). Research on upgrade path of vocational schools' personnel training quality under connotative development perspective. Journal of Luohe Vocational Technology College, 15(03), 17-19.

Marsh, H. W. (1987). Students' Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research. International Journal of Educational Research, 11(3), 253-388.

McNamara, T. (2000). Language testing. Oxford University Press.

Medgyes, P., & Nikolov, M. (2010). Curriculum development in foreign language education: The interface between political and professional decisions. in B. Kaplan, R. (2nd Ed.). The Oxford Handbook of Applied Linguistics. PP.263-274.

Nouraey, P., Al-Badi, A., Riasati, M. J., & Maata, R. L. (2020). Educational Program and Curriculum Evaluation Models: A Mini Systematic Review of the Recent Trends. Universal J Educ Res, 8(9), 4048-4055.

O'Leary, Z. (2014). The essential guide to doing your research project. 2nd ed. Thousand Oaks, CA: SAGE Publications.

Phillips, J. J. (2012). Return on investment in training and performance improvement programs. Routledge.

Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2019). Evaluation: A systematic approach. 8th ed. Sage Publications.

Stufflebeam, D. L. (2003). The CIPP Model for Evaluation. In Kellaghan, T., & Stufflebeam, D. L. (Eds.) International Handbook of Educational Evaluation. Springer, Dordrecht. pp. 31-62.