Implementation of Evaluation of Kirkpatrick Model in Statistics Course Based on Case Method at the Department of Building Engineering

Zulkifli Matondang[1*], Harun Sitompul[1], Kinanti Wijaya[1]

[1] Lecturer of the Department of Building Engineering Education, Universitas Negeri Medan, Medan, Indonesia.

Abstract: This study aims to obtain an overview of the implementation of Kirkpatrick's evaluation in case-based statistics courses based on Case Method at the Department of Building Engineering. Case-based lectures train students to solve problems in the field of statistics by applying higher-order thinking such as: finding solutions to problems, thinking critically, logically, reflectively, metacognitively and creatively. The evaluation of the Kirkpatrick model is carried out through four levels, and each level is assessed using a different tool. The implementation of Kirkpatrick's evaluation consists of 4 levels, namely: reaction, learning, behavior, and impact. This research method is descriptive qualitative research. The research sample is 56 students majoring in Engineering Building Engineering, State University of Medan. The research data was collected using a questionnaire, assessment and documentation. The questionnaire used has gone through reliability testing with Crombach's Alpha. The assessment was conducted to obtain student performance in the form of learning processes and behavior change. Documentation is obtained through the resulting lecture products. The results showed that Kirkpatrick's evaluation model in statistics course lectures was very effective in revealing the quality of learning. At the reaction evaluation level, students gave a positive response with satisfactory qualifications to case-based statistics lectures. The results of the evaluation at the learning level obtained the difference in learning outcomes on students' thinking abilities of 7.59 with a significance of 0.010. The results of the evaluation at the behavioral level, students have various competencies obtained such as: instrument preparation, instrument analysis, research data processing, SPSS application and interpretation of SPSS analysis results. The results of the evaluation at the impact level are that students are able to compile instruments, obtain and process data used to compose a final project. Based on the criteria for the success of Kirkpatrick's evaluation performance, the implementation of case-based statistics lectures was very successful. Kirkpatrick's evaluation model is able to describe evaluation results that cover all aspects of lectures so that it is worthy of being recommended as an evaluation model for lectures.

Keywords: Evaluation; Kirkpatrick model; Lectures; Performance assessment

Introduction

The Statistics course is one of the compulsory subjects in the Department of Building Engineering Education, Faculty of Engineering, Medan State University. The statistics course aims to make students familiar with concepts and data analysis in conducting research to complete their final project. Understanding the concept and analysis of research data, it is necessary for students to be able to think at a high level in statistics

lectures. In order to train students' higher-order thinking skills, an educator/lecturer must be able to present learning material in a simple way and in accordance with the problems that exist in the environment around students, especially in the field of statistics. In the statistics course, the learning method used is a case-based method.

The case method is a way of presenting subject matter using a case study flow. The indicators of the case method are first, the emphasis on certain cases; second,

to discuss cases intensively; third, emphasis on the case in detail (Laili et al., 2023). The case-based learning method is learning to solve problems thereby increasing critical thinking for problem solvers, communication skills, learning, and creativity (Rosidah et al., 2021). The case learning method can improve critical thinking skills so that it can solve cases of various calculations in statistics and is in line with the commitment of Higher Education Key Performance Indicators (IKU) in carrying out Case Based Method learning (Sofia et al., 2023).

In this case method-based statistics course, students are assigned to find solutions or problems. The case method in learning is used in developing critical thinking and collaboration skills (Anggraeni, 2020). With collaborative discussions, case-based learning methods can enhance study group interaction, decision making and problem solving. Given the complex characteristics of courses, we need a form of learning program evaluation that can reveal all abilities/potentials possessed by students, especially those related to statistics courses with case method-based learning. Evaluation has an important role in the running of a program. Without evaluation, educators do not know whether the program is successful or not so that follow-up can be carried out (Dewi et al., 2018).

Evaluation is the determination of the value of a thing, which can include information used to decide the value of the success of a program, product, procedure, goal or benefit in the design of an alternative approach to maintain its specific goals (Almanasreh et al., 2019; Tempelaar, 2019; Wu et al., 2019). Evaluation is a process that can provide information that is used as material for consideration to determine the value and price results of the goals achieved, design, implementation (Hanifa, 2018; Mulyanto et al., 2020). Results from evaluation can be a recommendation in making decisions, helping to account for, and increasing understanding of phenomena (Almutairi et al., 2021; Febriana, 2021; Peralbo-Uzquiano et al., 2020).

Meanwhile, in the opinion of Dimiyati (Irwantoro et al., 2016) revealed learning evaluation is a process for determining the value of learning carried out through learning assessment and measurement activities. This is confirmed by Mahirah (2017) who states that measurement, assessment, and evaluation are hierarchical, meaning that these three activities in relation to the learning process cannot be separated from one another and in their implementation must be carried out sequentially. Evaluation is preceded by an assessment, while assessment is preceded by measurement. Measurement can be interpreted as an activity to compare the results of observations or information on the characteristics of an object (Izza et al., 2020). Comparison of observation with criteria is measurement, interpretation and description of evidence is judgment and assessing the value or

implications of behaviour is evaluation. Evaluation is an integral part of lecturing activities for Statistics courses, in order to improve the quality of lectures.

According to Riadi (2017) the purpose of evaluation in general is to obtain evidentiary data which will be an indication of the level of ability and success of students in achieving learning objectives after taking the learning process within a specified time period. In addition, the evaluation also aims to measure and assess the extent to which teaching effectiveness and teaching methods have been applied by educators, as well as learning activities carried out by students. Evaluation results can be used as a benchmark for the success or failure of activities, can be continued or stopped, and can be used as a basis for developing follow-up activities. Evaluation activities as part of lecture activities need to be optimized. Evaluation is not only based on learning outcomes, but also on the input, output, and quality of the learning process.

An evaluation model is needed as a reference capable of evaluating the quality and effectiveness of the implementation of a training program (Ritonga et al., 2019). One widely known evaluation model that is adapted to education and is widely used as a strategy or work guideline in implementing learning program evaluations is the Kirkpatrick model (Cahapay, 2021). Previous research findings state that Kirkpatrick's model can be used as a training evaluation (Sakthi et al., 2021). Other findings also state that Kirkpatrick's model is effective in evaluating a program (Bijani et al., 2018; Buriak et al., 2019).

Kirkpatrick's model can be applied in various contexts, especially in social science evaluation to see high performance achievements in evaluating training. The results of Alsalamah and Callinan's research (2021) on the Kirkpatrick model show that research using this model is an active and developing area and effective in meeting the need to evaluate training activities. According to Kirkpatrick in Mahmoodi et al. (2019), there are three main reasons for evaluating using the Kirkpatrick evaluation model, namely improving the program, increasing the transfer of learning to behavior, then to results, and demonstrating the value of training to the organization. When the evaluation program has gone through changes, the evaluation can be used as an ongoing evaluation of an educational program that can guarantee its effectiveness.

The Kirkpatrick model is an effective training evaluation model based on four levels of evaluation, namely reaction, learning, behavior, and results (Bhatia et al., 2021; Hsu et al., 2021; Savul et al., 2021; Zielinska-Tomczak et al., 2021). According to Gracieux (2019) level 1 is Reaction, which assesses the extent to which participants find learning experiences enjoyable. Assessing how involved students are, how students contribute, and how students respond helps evaluators

to recognize how well students understand the training program. In this case, lecture activities are considered effective if the learning process is fun and satisfies students so that they are interested and motivated to learn and practice. In other words students will be motivated if the lecture process is running and satisfying for students which in the end will cause a pleasant reaction from students. Level 2 is Learning, which describes changes in knowledge, skills, and attitudes of participants. Evaluation of learning refers to the process and results of learning activities. The results of learning activities can be seen from the domains: knowledge, attitudes and skills. Students can be said to have learned if they have experienced changes in attitude, increased knowledge and increased skills. The quality of learning will determine student success. Its quality depends on how to present the material to be studied; how the teacher activates students, how the teacher uses reinforcement, how the teacher activates students to participate and feel involved in the teaching and learning process, how the teacher provides information to students about their success. Level 3 is Behavior, which indicates whether participants change their behavior and apply what they have learned. The criterion for the success of behavior evaluation is that the change in attitude that occurs after attending lectures will be carried out outside the lecture, so that this behavior assessment is more external in nature. Level 4 is Result, which measures the occurrence of the desired result. In learning activities this evaluation model leads to the final results obtained by students. Evaluation is carried out to obtain an overview or information about student learning outcomes and processes, as well as to find out the difficulties that arise during the teaching and learning process. Evaluation is also closely related to process assessment, both learning outcome assessment and process assessment.

Evaluation of Kirkpatrick's model can provide sufficient information to identify areas that require improvement, such as student responses to training programs, student understanding of training materials, and changes in student behavior after participating in training programs (Hidayat et al., 2023). By using the evaluation of the Kirkpatrick model, lecturers can thoroughly assess the effectiveness of learning in statistics courses based on the case method and make the necessary improvements to improve the quality of learning.

This study aims to obtain an overview of the implementation of Kirkpatrick's evaluation in case-based statistics courses. Case-based lectures train students to solve problems in statistics by applying higher order thinking. Evaluation of learning will affect whether or not learning objectives are achieved. Thus, evaluation activities are very important to measure the extent of student success in the teaching and learning process.

Method

This research is an evaluation method with a combination of qualitative and quantitative approaches. A combination approach is used to collect in-depth data and produce more comprehensive facts. The Kirkpatrick model evaluation research was carried out in statistics lecture activities for students majoring in Building Engineering Education, Faculty of Engineering, Unimed. The implementation of the Kirkpatrick model evaluation consists of four levels and each level is measured separately. The first level is in the form of evaluating reactions about student satisfaction with the process and implementation of lectures with the use of facilities. The second level is learning evaluation in the form of an assessment of lecture performance in terms of knowledge, attitudes and skills. Behavioral evaluation is the third level that is carried out to obtain an overview of student behavior after going through learning activities. The fourth level is more emphasized on the final result (student learning outcomes). The problem in the research is how to implement the Kirkpatrick model evaluation in the implementation of statistics course lectures with the case method.

To obtain data, at each level an assessment is used with its own instrument, in the form of questionnaires and tests. The instrument used has been tested to obtain validity and reliability. The focus of the evaluation is to obtain complete, in-depth data, and provide the right answer to the problem under study. The evaluation research procedure was carried out in a preparatory stage including the preparation of instruments for each level, the stages of punishment for the implementation of learning, as well as the stages of data collection and analysis. The research subjects were students of the Department of Building Engineering Education who took the Statistics course as many as 56 people consisting of 2 classes.

Data analysis is directed to describe the condition holistically. The collection of learning evaluation data is carried out through performance assessments, by making an assessment rubric. Rubric or assessment criteria is a description of the dimensions to decide student performance, value scales and standards to describe performance. The type of rubric used in the performance evaluation is divided into two forms, namely holistic and analytic rubrics. The holistic rubric is to describe the quality of the performance while the analytic rubric provides scores for the task components. The scoring rubric used in this study is a holistic rubric as shown in Table 1.

Table 1. Rubric of General Criteria Level Assessment Criteria

Category	Meaning (Rublic)
4 (Very satisfy)	Demonstrate a greater understanding of concepts, Complete lecture information, using appropriate strategies and assessment systems, easy access to learning facilities and resources, exceeding the desired problem solving.
3 (Satisfying)	Demonstrates understanding of concepts, uses appropriate strategies, calculations are mostly correct, explanations are effective, fulfills all problem solving
2 (Good enough)	Demonstrates understanding of most concepts, does not use appropriate strategies, calculations are mostly correct, explanations are satisfactory, fulfills most of the desired problem solving
1 (Not satisfactory)	Shows little understanding of concepts, does not use appropriate strategies, calculations are not correct, explanations are not satisfactory, does not meet the desired problem solving.

Result and Discussion

The implementation of the evaluation of the Kirkpatrick model is carried out in the learning process in the statistics course. The learning time is adjusted to the lecture schedule and the number of course credits. The results of data analysis of case method learning activities using Kirkpatrick's model evaluation in statistics courses are described in four evaluation aspects. Each aspect of the evaluation is analyzed to obtain information about the ability of students majoring in building engineering education in the statistics course. Data obtained at each level of evaluation, using different assessment tools. The results for each level are as follows.

Evaluating Reaction

The first evaluation of Kirkpatrick's model is Reaction. The learning program or lecture process is considered successful if students are satisfied with the strategy for delivering the material, the learning media used, and the assessment used. It was concluded that the success of the learning process was inseparable from the interest, attention, and motivation of students in following the course of the lecture. Students learn better when they react positively to the learning environment. Student satisfaction with the learning process carried out was assessed from several aspects, namely the description of the material provided, the available facilities, the strategy for delivering the material by the lecturer, and the available learning media. The description of the level of student satisfaction with the learning process carried out is as shown in Table 2.

Table 2. Student Satisfaction Levels with Statistics Lectures

Satisfaction Indicator	Average
Complete Information About Lectures	3.25
Easy Access To Learning Facilities And Resources	3.52
The Suitability Of The Material With The Course Planning	3.17
Determination Of The Process Or Method Of Teaching	2.91
Professionalism And Teaching Competence	3.05
Responsibility Of Teachers And Employees	2.82
Suitability Of The Scoring System	3.27
Accuracy Of The Process With The Results Obtained	3.18
Average Score	3.15

The results of the evaluation of reactions as shown in Table 2, show that students majoring in building engineering education at FT Unimed who took the statistics course gave a positive reaction to the learning process carried out. The data shows that the highest score of student satisfaction is on the indicator of the ease of access to facilities and learning resources with a score of 3.52 out of a maximum score of 4. While the lowest score is on the indicator of the responsibility of teaching staff and employees with a score of 2.27 out of a score of 4. The second lowest score is the indicator of the accuracy of the process and lecture method with a score of 2.91 out of a maximum score of 4. On average, student satisfaction with the implementation of case-based statistics lectures with a score of 3.15 is in the satisfactory category.

The results of the evaluation of this reaction indicate that the indicators of the responsibility of teaching staff and employees as well as indicators of the accuracy of the learning process and methods need attention to improve the quality of statistics course lectures in the Department of Engineering Education, Faculty of Engineering, Unimed. The results of this reaction evaluation reveal that the responsibility of teachers and employees needs to be improved. This shows that students want the teaching staff (lecturers) and staff to be responsive to serve students in carrying out better learning activities. This can be done in the service of providing learning facilities or facilities in the Department of Building Engineering Education such as providing: access to learning resources in the library, infocus, lecturer presentation screens, electrical cables and stationery on the withboard. Meanwhile, services provided by educators include: opportunities to discuss outside of learning with students and provide access to various learning resources relevant to statistical material. The results of this evaluation become a reference for making improvements or improving the quality of the lecture service process by lecturers and employees.

Evaluating Learning

The second level of evaluation is the Kirkpatrick learning evaluation model. Students who attend lectures are assessed for their abilities and knowledge through performance assessments. Performance assessment is carried out in order to find out all the potential of students. This second level research describes the aspects of the overall assessment and clear criteria. Performance appraisal is part of an alternative assessment that is realistic and can provide meaningful feedback in developing student potential.

Alternative assessment is needed to assess the dimensions of the process and student learning outcomes that are not obtained through giving or assessment through tests. Experts suggest that the use of traditional tests (paper and pencil tests) in the learning process has not been able to assess all aspects of learning. The test can only assess scientific knowledge, not able to develop the potential possessed by students, especially those related to course competencies.

According to Nurhayati (2018) the steps taken in the evaluation of the second level Kickpatrick model, namely evaluating related: (a) increasing knowledge, skills, and attitude changes before and after lectures; (b) attitude to use the agreed test indicators; (c) knowledge using pretest and posttest; and (d) skills using performance tests. Actions taken to confirm the results of the evaluation at the reaction level include: whether the teacher is less communicative in delivering the material, whether the learning strategies are carried out in accordance with student expectations, or there are other factors at the first level that might cause students to experience demotivation in learning.

Table 3. Average Pretest and Posttest Students of the Department of Building Engineering Education

	N	Minimum	Maximum	Mean	Std. Deviation
Pretest	56	63	84	72.16	5.055
Postest	56	65	90	79.75	5.949
Valid N (listwise)	56				

The second level evaluation activity in this paper is to measure knowledge through pretest and posttest. The results of the average pretest and posttest of students majoring in Building Engineering Education who take statistics courses are presented as follows table 3.

Based on Table 3 shows that the pretest score of students majoring in Engineering Education who took 56 statistics courses was obtained an average of 72.16 with a standard deviation of 5.055. Furthermore, after learning with the case method approach, students were carried out post-test. The results of the posttest score obtained an average of 79.75 with a standard deviation of 5.949. Statistically there is a difference in the average score of the posttest score with the average score of the pretest score of 7.59. The score difference between the

average posttest and pretest scores from students majoring in building engineering education who took statistics courses was relatively large. Then to find out the significance of the difference in the average score between the pretest and the posttest, an analysis of variance (ANAVA) test was carried out with the help of SPSS. The results of the analysis are presented in the following table 4.

Table 4. Results of Analysis of Variance from Pretest and Posttest Scores

	Sum of Squares	df	Mean Square	F	Sig
Between Groups	309.185	1	309.185	10.197	.002
Within Group	1637.315	54	30.321		
Total	1946.500	55			

Table 4 shows that based on the analysis of variance from the pretest and posttest scores of students majoring in Building Engineering Education who took the statistics course, the calculated F value was 10.197 with a sig score. of 0.002. These results indicate that there is a significant or significant difference between the average posttest score and the average pretest score. The average posttest score was 7.59 greater than the average pretest score. This shows that the learning method with a case approach in the statistics course is quite significant in increasing students' knowledge.

Assessment on the second level evaluation for aspects of attitudes and skills is carried out qualitatively through observation during lectures. In general, students who take statistics courses with the case method approach make the students more active and cooperative in solving the given cases. The students tried to find various learning resources to be able to solve the cases given, and between students more discussions and exchanged information obtained. Thus, students' understanding of the concepts of various statistical materials is better and improved. It can be concluded that student achievement in mastering statistical material is increasing. This reveals that the learning evaluation for statistics courses in the cognitive, affective and psychomotor domains is quite effective through the case study method.

Evaluating Behavior

The third level of evaluation of the Kirkpatrick model in the statistics course learning program is very important to measure students' ability to apply or apply statistical concepts in modeling related to everyday life phenomena. In learning statistics courses, students are given the opportunity to make a data analysis and interpretation. Data analysis is carried out in the form of applying statistical concepts and interpretations that can be applied in everyday life. The truth of the results of the analysis and interpretation carried out by students

becomes information on the extent of student mastery of the statistical concepts that have been studied.

Behavior evaluation to see how far changes in student behavior occur after attending lectures. So a behavioral evaluation is carried out to identify the extent to which the material in lectures can be applied in life or help complete the final project. The purpose of conducting an evaluation at the behavior stage is to measure changes in student behavior that arise as a result of attending lectures. To implement behavior change, according to Kirkpatrick in Ritonga (2019) there are four conditions needed, namely a person must have the desire to change, a person must know what to do and how to do it, a person must work in a decent work environment, and a person must be respected for He changes. Lecture programs can provide conditions that support behavior change in accordance with lecture objectives by providing material related to knowledge, skills, or attitudes. Behavior that can be changed through learning statistics courses using the case method for students majoring in building engineering education, such as behavior or abilities in the fields of: compiling instruments, testing instruments, obtaining data from the field, sorting data, analyzing data manually or with SPSS. In general, changes in students' behavior or competence in taking statistics courses are quite good through learning using the case method.

Evaluating Result

The fourth level of Kirkpatrick's model is Result. The implementation of learning activities, of course, aims to get good results, such as improving quality, productivity, or safety levels. Outcome evaluation is defined as an end result that occurs as a result of students attending lectures. The steps in evaluating the results are: (1) Conduct a level-3 evaluation, namely behavioral evaluation. (2) Give time to see the impact that appears or is achieved. There is no specific time to evaluate the results, so that in determining the time of the evaluation, the various factors involved must be considered. (3) Can be done by survey method using a questionnaire. or interviews with students and lecturers or student final project supervisors. (4) Take measurements, both before and after the implementation of learning if possible. (5) Perform re-evaluation at the appropriate time at the appropriate time. (6) Consideration of costs incurred with the results obtained. (7) Can use secondary data, such as student competency data in other subjects, and other data that supports survey results in analyzing results. In implementing the four-level evaluation model, it must be done sequentially, because each level is important and has an impact on the next level.

The last level of the evaluation of the Kirkpatrick model in the statistics course learning program is emphasized on the final results of student academic activities. Student success can be seen at the third level of the Kirkpatrick evaluation model. The assessment criteria at this level are a continuation of the first level assessment, namely the use of statistical material concepts. Students obtain data and perform analysis and interpretation of data that can be used in everyday life. The results that have been obtained through the analysis are carried out through data processing using SPSS software. With the ability to analyze through computation and data interpretation, students can process data and compile reports in completing the final project.

Kirkpatrick's evaluation model makes it easy for lecturers to conduct a thorough evaluation of the potential possessed by students. Statistics courses, which have the characteristics of analysis and computational applications, require an evaluation model that is able to explore student performance as a whole. The evaluation of the Kirkpatrick model which is applied to the learning of statistics courses with four levels of evaluation provides an overview of the abilities possessed by students. The level of student mastery of a subject matter or subject matter can be determined by several aspects, with varying scoring systems. A varied scoring system can reveal all the potential or abilities of students in each indicator of the course. The assessment rubric used at each level in the evaluation of the Kirkpatrick model can motivate students to reveal all that is known about competencies in statistics courses.

The design of the learning process should pay attention to the evaluation model used and the form of assessment used, to get a comprehensive picture of the competencies possessed by students. An assessment process that is able to reveal all the potential of students, can create an active and efficient learning atmosphere. This is because student activities in learning activities are more focused on the ability to solve problems or reveal all that is known about existing courses. In relation to statistics courses that involve the use of computers in the form of SPSS software, project or performance assessments are needed. An assessment system that only prioritizes written tests, it will be difficult to measure student success against the competencies of statistics courses.

In the learning process carried out in tertiary institutions, the lecturer acts as an evaluator whose function is to measure the achievement of the objectives that have been formulated and the material that has been determined. Lecturers as assessors of learning outcomes or student learning processes, should be able to follow continuously on student achievements. The information obtained becomes feedback on the learning process carried out. The results of the feedback become a reference in improving and improving the learning process carried out by lecturers, thus optimal and effective learning will be achieved.

The evaluation process carried out by lecturers for statistics courses can be carried out at the end or during the teaching and learning process. Lecturers giving courses can evaluate by taking into account the criteria for completeness obtained by students. Implementation of evaluation by lecturers should be carried out continuously in each learning process. If the evaluation process is only carried out at the end of learning, then the process carried out during learning activities cannot be described in the final results of students. The evaluation carried out should be able to provide information about the achievement of predetermined goals. The evaluation process carried out must be able to reveal the intelligence possessed by students, both cognitive, effective and psychomotor.

Student responses to the evaluation of the Kirkpatrick model and the assessment rubric used are very good, this is indicated by the results of student work independently and in groups. In the academic evaluation of students also showed good results. The success of the learning carried out is largely determined by the responses and results obtained by students. In conducting an evaluation, the evaluator should pay attention to the aspects to be achieved, the target students and the assessment rubric.

The theoretical implication of the results of this study is that the evaluation of the Kirkpatrick model provides information on how to assess student abilities as a whole. The practical implication is that the evaluation of the Kirkpatrick model is an alternative for lecturers to get an overview of students' abilities in statistics or other subjects. The results of this study indicate that by using Kirkpatrick's evaluation model in the learning process of statistics courses at the Department of Building Engineering, FT Unimed, student learning activities are very effective and able to explore all student abilities. This research has an impact on the evaluation of learning in general for other courses with the same characteristics. Assessment in each learning process is carried out in various ways to get an optimal picture of the potential in students, especially those related to certain subjects. This is done to improve student learning outcomes. Limitations this research was conducted by following the stages or levels in the evaluation of the Kirkpatrick model, but there are some limitations of the study including, the number of samples or research classes consisting of two classes. The process of observation carried out in every aspect of each rubric and the number of students who were sampled became one of the limiting factors in this study.

Conclusion

The results of the research and discussion through Implementation of Evaluation of Kirkpatrick Model in Statistics Course Based on Case Method at the Department of Building Engineering can be concluded, namely: a) Statistics lectures are very effective in revealing students abilities, especially about performance and competence in statistics course; b) students gave a positive response to case-based course lectures; c) students experienced increased thinking skills in applying statistical concepts in everyday life with an average score difference of 7.59; d) through case-based statistics lectures, students have various competencies such as: preparation and analysis of instruments, collection and processing of research data, data analysis with the SPSS application and interpretation of the output. e) impact evaluation results, namely students are able to develop instruments and obtain research data and analyze them so that they can compiling a thesis as a final project in completing the study.

Author Contributions
ZM conducted studies and participated in data collection, statistical analysis, writing of manuscripts. HS participated in the research design and conducted data interpretation. KW drafted the study, and participated in its design and coordination and assisted in drafting the manuscript. All authors contributed from start to finish and approved the final manuscript.

Conflicts of Interest
All authors in this study declare that they have no competing interests to all parties.

References

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, *15*(2), 214–221. https://doi.org/10.1016/j.sapharm.2018.03.066

Almutairi, T. S., & Shraid, N. S. (2021). Teacher evaluation by different internal evaluators: Head of departments, teachers themselves, peers and students. *International Journal of Evaluation and Research in Education (IJERE)*, *10*(2), 588. https://doi.org/10.11591/ijere.v10i2.20838

Alsalamah, A., & Callinan, C. (2021). The Kirkpatrick model for training evaluation: bibliometric analysis after 60 years (1959–2020). *Industrial and*

Commercial Training, *54*(1), 36–63. https://doi.org/10.1108/ICT-12-2020-0115

Anggraeni, L. (2020). Penerapan Metode Studi Kasus Dalam Upaya Meningkatkan Kemampuan Berpikir Kritis Mahasiswa pada Matakuliah Hubungan Internasional. *Media Komunikasi FPIPS*, *10*(2), 23. https://doi.org/10.23887/mkfis.v10i2.462.

Bhatia, M., Stewart, A. E., Wallace, A., Kumar, A., & Malhotra, A. (2021). Evaluation of an In-Situ Neonatal Resuscitation Simulation Program Using the New World Kirkpatrick Model. *Clinical Simulation in Nursing*, *50*, 27–37. https://doi.org/10.1016/j.ecns.2020.09.006

Bijani, M., Rostami, K., Momennasab, M., & Yektatalab, S. (2018). Evaluating the Effectiveness of a Continuing Education Program for Prevention of Occupational Exposure to Needle Stick Injuriesin Nursing Staff Based on Kirkpatrick's Model. *Journal of the National Medical Association*, *110*(5). https://doi.org/10.1016/j.jnma.2017.11.002.

Buriak, S. E., & Ayars, C. L. (2019). Evaluation of a drug and alcohol safety education program in aviation using interrupted time series and the Kirkpatrick framework. *Evaluation and Program Planning*, *73*, 62–70. https://doi.org/10.1016/j.evalprogplan.2018.11.003

Cahapay, M. (2021). Kirkpatrick model: Its limitations as used in higher education evaluation. *International Journal of Assessment Tools in Education*, *8*(1), 135–144. https://doi.org/10.21449/ijate.856143

Dewi, L. R., & Kartowagiran, B. (2018). An evaluation of internship program by using Kirkpatrick evaluation model. *REID (Research and Evaluation in Education)*, *4*(2), 155–163. https://doi.org/10.21831/reid.v4i2.22495

Febriana, R. (2021). *Evaluasi pembelajaran*. Bumi Aksara.

Gracieux, F. Y. (2019). The Asian Sarcoma Consortium Sarcoma Preceptorship Program: A Program Evaluation Study Utilizing the Kirkpatrick Model (Level 1 and 2). *Annals of Oncology*, *30*(9). https://doi.org/10.1093/annonc/mdz433.005

Hanifa, R. (2018). EFL Published Materials: An Evaluation of English Textbooks for Junior High School in Indonesia. *Advances InLanguage and Literary Studies*, *9*(2), 166. https://doi.org/10.7575/aiac.alls.v.9n.2p.166

Hidayat, R., Alam, B. P., Lutvaidah, U., & Santosa, P. P. P. (2023). Implementasi Evaluasi Model Kirkpatrick Pada Program Pelatihan Penggunaan Eyd V SMK Broadcasting Mahardika. *Abdi Jurnal Publikasi*, *1*(5), 436–441. Retrieved from https://jurnal.portalpublikasi.id/index.php/AJP/article/view/173

Hsu, W., & Chen, P.-W. (2021). The influences of service quality and individual characteristics on vocational training effectiveness. *Sustainability*, *13*(23), 13207. https://doi.org/10.3390/su132313207

Irwantoro, N., & Suryana, Y. (2016). *Kompetensi pedagogik*. Genta Group Production.

Izza, A. Z., Falah, M., & Susilawati, S. (2020). Studi literatur: problematika evaluasi pembelajaran dalam mencapai tujuan pendidikan di era merdeka belajar. *Konferensi Ilmiah Pendidikan Universitas Pekalongan 2020*, *1*, 10–15. Retrieved from https://proceeding.unikal.ac.id/index.php/kip

Laili, S. N., Sudarti, S., & Supriadi, B. (2023). The Effect of Physics Learning Based on Case Method on Skills in Inferences Problems Temperature and Health. *Journal of Science and Science Education*, *4*(1), 50–54. https://doi.org/10.29303/jossed.v4i1.2556

Mahirah. (2017). Evaluasi Belajar Peserta Didik (Siswa). *Idaarah: Jurnal Manajemen Pendidikan*, *1*(2). https://doi.org/10.24252/idaarah.v1i2.4269

Mahmoodi, M., Rashtchi, M., & Abbasian, G. R. (2019). Evaluation of in-service teacher training program in iran: Focus on the Kirkpatrick model. *Education and Self Development*, *14*(4), 20–38. https://doi.org/10.26907/esd14.4.03.

Mulyanto, B. S., Sadono, T., Koeswanti, H. D., Dasar, S., Wonodoyo, N., Tengah, J., Kristen, U., Wacana, S., & Tengah, J. (2020). Evaluation of Critical Thinking Ability with Discovery Lerning Using Blended Learning Approach in Primary School. *Journal of Educational Research and Evaluation*, *9*(2), 78–84. https://doi.org/10.15294/jere.v9i2.46135.

Nurhayati, Y. (2018). Penerapan Model Kirkpatrick untuk Evaluasi Program Diklat Subtantif Materi Perencanaan Pembelajaran di Wilayah Kerja Provinsi Kepulauan Riau. *Andragogi Jurnal Diklat Teknis*, *VI*(2), 170–187. https://doi.org/10.36052/andragogi.v6i2.63

Peralbo-Uzquiano, M., Fernández-Abella, R., Durán-Bouza, M., Brenlla-Blanco, J.-C., & Cotos-Yáñez, J.-M. (2020). Evaluation of the effects of a virtual intervention programme on cognitive flexibility, inhibitory control and basic math skills in childhood education. *Computers & Education*, *159*, 104006. https://doi.org/10.1016/j.compedu.2020.104006

Riadi, A. (2017). Problematika Sistem Evaluasi Pembelajaran. *Ittihad Jurnal Kopertais Wilayah XI Kalimantan*, *15*(27), 11. Retrieved from http://jurnal.uin-antasari.ac.id/index.php/ittihad/article/viewFile/1593/1162

Ritonga, R., Saepudin, A., & Wahyudin, U. (2019). Penerapan model evaluasi kirkpatrick empat level dalam mengevaluasi program Diklat di Balai Besar Pelatihan Pertanian (BBPP) Lembang. *Jurnal*

Pendidikan Nonformal, *14*(1), 12–21. https://doi.org/10.17977/um041v14i1p12-21

Rosidah, C., & Pramulia, P. (2021). Team Based Project dan Case Method Sebagai Strategi Pengembangan Keterampilan Mengembangkan Pembelajaran Mahasiswa. *MENDIDIK: Jurnal Kajian Pendidikan Dan Pengajaran*, *7*(2), 245–251. https://doi.org/10.30653/003.202172.196

Sakthi, T., & Moshi, A. A. M. (2021). Effectiveness measurement study on entrepreneurship awareness camp for technical education students using Kirkpatrick's model. *Materials Today: Proceedings*, *45*(2), 1609–1613. https://doi.org/10.1016/j.matpr.2020.08.419

Savul, S., Ikram, A., Khan, M. A., & Khan, M. A. (2021). Evaluation of Infection Prevention and Control Training Workshops using Kirkpatrick's Model. *International Journal of Infectious Diseases*, *112*, 76–80. https://doi.org/10.1016/j.ijid.2021.09.005

Sofia, N., Ritonga, M., Arita, S., Syahrul, S., Shalihah, M. A., Dewi, I. P., & Sofya, R. (2023). Peningkatan Keterampilan Berpikir Kritis Melalui Pengembangan E-Modul Perpajakan Berbasis Case Method. *Jurnal Ecogen*, *6*(1), 54–63. https://doi.org/10.24036/jmpe.v6i1.14185

Tempelaar, D. (2019). Assessment & Evaluation in Higher Education Supporting the Less-Adaptive Student: TheRole of Learning Analytics, Formative Assessment and Blended lLarning. *Assessment and Evaluation in Higher Education*, *45*(4). https://doi.org/10.1080/02602938.2019.1677855.

Wu, W.-H., Kao, H.-Y., Wu, S.-H., & Wei, C.-W. (2019). Development and Evaluation of Affective Domain Using Student's Feedback in Entrepreneurial Massive Open Online Courses. *Frontiers in Psychology*, *10*(May). https://doi.org/10.3389/fpsyg.2019.01109

Zielinska-Tomczak, L., Przymuszała, P., Tomczak, S., Krzyśko-Pieczka, I., Marciniak, R., & Cerbin-Koczorowska, M. (2021). How Do Dieticians on Instagram Teach? The Potential of the Kirkpatrick Model in the Evaluation of the Effectiveness of Nutritional Education in Social Media. *Nutrients*, *13*(6). https://doi.org/10.3390/nu13062005