Journal of Educational Research and Evaluation
Volume 6, Issue 3, 2022, pp. 688-695
P-ISSN: 2597-422x E-ISSN: 2549-2675
Open Access: https://doi.org/10.23887/jere.v6i4.54695

Summative Evaluation of ICT-Based Learning Media

Ni Made Sri Mertasari[1*]

[1] Program Studi Pendidikan Matematika, Universitas Pendidikan Ganesha, Indonesia

ARTICLE INFO

Article history:
Received August 04, 2022
Revised August 05, 2022
Accepted November 14, 2022
Available online November 25, 2022

Kata Kunci :
Formatif, evaluasi, berbasis TIK, Media Pembelajaran

Keywords:
Formative, evaluation, ICT-based, Learning Media

ABSTRAK

Mahasiswa calon guru telah banyak terlibat dalam pengembangan media pembelajaran berbasis TIK. Berbagai bentuk media telah diproduksi, seperti multimedia, hypermedia, web, video, animasi, atau presentasi. Sampai saat ini belum semua media yang dikembangkan mengalami evaluasi sumatif melainkan hanya sampai pada evaluasi formatif. Salah satu penyebabnya adalah kurangnya referensi. Penelitian ini bertujuan untuk merumuskan model evaluasi sumatif terhadap media pembelajaran berbasis TIK yang dikembangkan oleh mahasiswa calon guru dengan menggunakan pendekatan deskriptif kuantitatif dan kualitatif. Penelitian diawali dengan observasi terhadap laporan penelitian siswa untuk mengkaji model evaluasi sumatif yang digunakan. Informasi tersebut kemudian dibahas dalam diskusi kelompok terarah yang melibatkan pakar dan praktisi terkait. Studi ini menemukan bahwa evaluasi sumatif terhadap media pembelajaran berbasis TIK dapat dilakukan melalui survei kuasi eksperimen atau studi kasus. Metode kuasi eksperimen yang dapat diterapkan adalah nonequivalent control group design, factorial design, counterbalanced design, dan time series design. Format tersebut diperoleh terutama dengan mempertimbangkan waktu yang terbatas, tempat ujian, dan kesempatan untuk terlibat di dalam kelas. Baik eksperimen semu, survei, maupun studi kasus dapat dilakukan oleh siswa bekerjasama dengan guru terkait saat siswa melakukan praktikum mengajar di sekolah dan kemudian dimatangkan saat siswa melakukan penelitian.

ABSTRACT

Prospective teacher students have been heavily involved in developing ICT-based learning media. Various forms of media have been produced, such as multimedia, hypermedia, web, video, animation, or presentation. Until now, not all media developed have undergone a summative evaluation but only arrived at a formative evaluation. One of the reasons is the lack of references. This study aims to formulate a summative evaluation model of ICT-based learning media developed by prospective teacher students using quantitative and qualitative descriptive approaches. The study began with observations of student research reports to examine the summative evaluation model used. The information is then discussed in a focus group discussion involving relevant experts and practitioners. The study found that the summative evaluation of ICT-based learning media could be done through a quasi-experimental survey or case study. Quasi-experimental methods that can be applied are nonequivalent control group design, factorial design, counterbalanced design, and time series design. The format was obtained mainly considering limited time, a place for testing, and opportunities to be involved in the class. Either quasi-experiments, surveys, or case studies can be carried out by students in collaboration with relevant teachers when students do teaching practicum at school and then mature when students do research.

1. INTRODUCTION

Evaluation is the process of determining the value of a product or process (Scriven, 1992; Vedung, 2017). The term evaluation was originally known in education to assess objects (Patton, 1994). There are several types of evaluation in the learning process, including formative and summative evaluation. Formative evaluation is carried out by the teacher during the learning process to improve the quality of learning, while summative evaluation is carried out at the end of the learning program to see the success of the learning program (Popham, 1999; Spivey, 2007). Therefore, formative evaluation is evaluation as learning, while summative evaluation is evaluation for learning (Amin & Eng, 2003; Deeley, 2018; Earl, 2003). Adopting the term evaluation in the learning process, the development of learning media also recognizes the terms formative evaluation and summative evaluation (Chang, 2006; Nieveen & Folmer, 2013; Plomp & Nieveen, 2013). Formative evaluation is carried out during the development process to increase effectiveness, while summative evaluation is carried out at the end of development to determine the acceptability and usefulness of the developed learning media (Aldoobie, 2015; Drljača et al., 2017).

*Corresponding author.
E-mail addresses: srimertasari@undiksha.ac.id (Ni Made Sri Mertasari)

The learning media developed are expected to meet the target criteria, such as effectively improving learning achievement, efficiently utilizing facilities, matching the characteristics of students, meeting the satisfaction of educators and students, and other set targets. Therefore, even though it has undergone formative evaluation, the resulting learning media must go through a summative evaluation. Formative evaluation has traditionally focused on the validity, accuracy, and functioning of instructional media (Hattie et al., 2006; Weston, 2004), but only once the assessment of implementation problems (Weston, 2004). Therefore, after the media is developed, it is necessary to conduct a summative evaluation to obtain evidence of its effectiveness and find arguments supporting the decision to continue or stop the project (Nieveen & Folmer, 2013).

Summative evaluation focuses on the practicality and effectiveness of learning media (Plomp & Nieveen, 2013). Effective in the sense of bringing students to achieve the expected learning achievements. Furthermore, practical is interpreted as easy to use. In other words, the developed learning media makes it easy for students to use. Summative evaluation is also used to test instructional media to show evidence of the compatibility of technology with learning content (Mckenzie, 2006). The summative evaluation aims to get the big picture and assess ready-made learning media (Joyce, 2019). Therefore, evaluation is carried out only at the end of development, unlike formative evaluation, which is carried out at each stage of development. Summative evaluation can be divided into expert judgment and field trial (Chang, 2006). Expert judgment is intended to evaluate the suitability of the information presented with the instructions needed by students, the completeness and accuracy of the instructions, the accuracy of learning strategies, the usefulness of learning instructions, and student satisfaction. Meanwhile, the field trial is intended to test the effectiveness of the developed product. Two things are evaluated during the field trial: 1) the impact of teaching on student competence and 2) the attitudes of teachers and school management related to student performance, implementation feasibility, organizational culture fit, and cost suitability. Evaluation is done by implementing instructional products in learning practices. Of course, the instructional products developed have gone through formative to final evaluation.

Summative evaluation is used to provide an overview of the overall performance of a learning media. The summative evaluation process is often carried out by comparing it with previous versions or competing media (Mckenzie, 2006) or comparing it with the requirements or benchmarks set (Nieveen & Folmer, 2013). The results obtained from the summative evaluation describe the actual media performance, not limited to perceptions. Evaluation is carried out quantitatively or qualitatively to test the system's functionality. In addition, summative evaluation measures effectiveness and ease of use, so it must include the level of task completion, time to do the task, error rate, and overall user satisfaction. The final desired decision from the summative evaluation is a recommendation to continue or stop development (Nieveen & Folmer, 2013). Summative evaluation can be done through a quasi-experimental (Nieveen & Folmer, 2013; Wayne et al., 2008). Experiments are used to explain and compare (Plomp & Nieveen, 2013). The quasi-experiment was used to test the effect of intervention without randomization (Eraut, 1982; Fraenkel, Jack R., Wallen, 2009). Existing groups or classes are left because conditions do not allow individual randomization. For example, a class in a school that has been formed with certain considerations is unlikely to be changed again by forming a new class through individual randomization only for research purposes. Several research design models are included in the quasi-experimental design, such as the nonequivalent control group design, the time-series design, the counterbalanced, and the factorial design (LR et al., 2006).

Nonequivalent control group design is an experiment that compares the experimental group that received the intervention and the control group that did not receive the intervention, which group was not determined randomly (Gall et al., 2003; Gravetter, F. J., Forzano, 2009). If the experimental group is superior to the control group, it can be concluded that the intervention affects the subject. Time series design was carried out with a series of observations of one group before and after the intervention (Gall et al., 2003; Gravetter, F. J., Forzano, 2009). Observations before the intervention and after the intervention should be the same. In addition, the inter-observation interval before and after the intervention should also be the same. The discrepancy between the observation results after and before the intervention showed the influence of the intervention on the subject. A factorial design is any design involving two or more independent variables, at least one of which is manipulated to determine whether the effects of the independent variables can be generalized across all levels or whether the effects are specific to a particular level (Fraenkel, Jack R., Wallen, 2009; LR et al., 2006). The group of variables investigated in a factorial design is called a factor. The effect of each independent variable on the dependent variable is called the main effect. In contrast, the effect of the interaction between independent variables on the dependent variable is called the interaction effect (Gall et al., 2003). The advantage of the factorial design is the opportunity for researchers to observe two variables that influence and interact simultaneously. The counterbalance design is usually used when the study involves the whole group because it does not

allow randomization, and a pretest is also impossible (LR et al., 2006). In a counterbalance design, all groups received interventions differently (Fraenkel, Jack R., Wallen, 2009; Gravetter, F. J., Forzano, 2009; Sarkies et al., 2019). In other words, the experimental and control interventions were applied to all groups in different orders. Counterbalancing allows researchers to control for the effects of confounding variables in the design (Allen, 2019; Zeelenberg & Pecher, 2015). Observations for all groups were made at the end of each intervention. Research decisions were made by comparing the mean observations for each inter-group intervention.

Case studies or surveys can also do the summative evaluation (Nieveen & Folmer, 2013). The case study includes an intensive study of one research unit, while the survey takes a population sample and uses a questionnaire as the main data collection tool (Scriven, 1992). Case studies were conducted on individuals to obtain a description of the individual (Gravetter, F. J., Forzano, 2009). An in-depth study of individuals is carried out in a case study, so research generally only involves a small group. Researchers intervened in a group of individuals within a specified period. Furthermore, the researchers conducted in-depth observations of the group of individuals to reveal the intervention's effect on individual characteristics. Several case studies apply longitudinal research by conducting observations, data collection, data analysis, and reporting systematically over a long period (Symonds & Symonds, 2016). If researchers want to study phenomena that arise from certain entities, case studies allow them to provide in-depth understanding (Heale & Twycross., 2018). Even though they are considered less thorough and less objective, case studies are still widely used, especially when other approaches are impossible (Rowley, 2002). The survey is also an alternative method of summative evaluation. The survey is a descriptive research method used to collect primary data based on verbal or written communication with a representative sample of individuals or respondents from the target population (Mathiyazhagan et al., 2010). Surveys can be used as a data collection method to test hypotheses or to answer questions about some topic or problem (Gay & Mills, 2012). Surveys are also often used to collect information about attitudes and behavior. For ethical or practical reasons, researchers are unlikely to intervene in one group and choose another group as a control (Navarro-Rivera & Kosmin, 2015). Current surveys can reflect various research objectives, sampling techniques, data collection instruments, and survey administration methods (Ponto, 2015).

2. METHODS

This research aims to get a formula for the summative evaluation method of learning media that prospective teacher students can use as a reference. Therefore, this study begins with a quantitative descriptive analysis to identify data patterns (Loeb et al., 2017) and then reveals the phenomenon and its characteristics (Nassaji, 2015). Student research reports on learning media development for the last five years were collected and observed to identify the summative evaluation method used. Observations were made by two experts who kept in touch to avoid data bias. A third expert used triangulation to ensure that the data had come from different perspectives to help validate the data (Wüthrich, 2004). Quantitative data are presented in tabular form and then interpreted. Observations also collect qualitative data about the reasons for choosing the summative evaluation method. The qualitative data obtained were checked by the relevant experts. Data is declared valid if it converges with expert assessments (Creswell, 2015; Miles et al., 2014; Springer, 2010). The first time, data simplification was done through data condensation to select keywords so that the data was more focused on the research objectives. The next step is to organize the data through the data display so that the information is well organized and concluding is easier. A visual display is performed to make the data set more explicit about assisting the inference process (Verdinelli & Scagnoli, 2013). If the information series is visible, the next step is drawing and verifying the conclusion.

If the initial conclusions are still considered weak, then the display data results are reviewed, and there is even an opportunity for re-analysis. It is also possible to return to condensation data if the conclusions are weak. The worst situation, namely re-observation, can occur if the data that has been collected is considered insufficient. The process of data condensation, data display, visual display, and drawing and verifying conclusions can be repeated as needed until the conclusions obtained are considered good. The conclusions are verified through theoretical triangulation, namely by interpreting based on the perspectives of several theories (Farquhar & Michels, 2016). Thus, the conclusions obtained are stronger because apart from being supported by empirical data, they are also supported by existing theories.

3. RESULT AND DISCUSSION

Results

Observations only found eighty-four research reports on developing information technology-based learning materials containing summative evaluations in the last five years. The rest research reports on the development of information technology-based learning materials made by prospective teacher students do not include summative evaluations. These findings still need to be able to conclude the summative evaluation model of information technology-based learning materials. The research sample was expanded by involving research reports that tested information technology-based learning materials previously developed by other researchers to achieve the research objectives. Observations identified the summative evaluation method used by prospective teacher students in developing ICT-based learning media, as listed in the Table 1.

Table 1. Frequency of Application of Summative Evaluation Method

Method	Frequency	Percentage
Quasi-experiment and Case Study	28	10.85
Quasi-experiment and Survey	52	20.16
Quasi experiment	118	45.74
Case study	19	7.36
Survey	41	15.89
Total	258	100.00

Most researchers applied quasi-experimental for summative evaluation, as many as 198 or 76.74%. The quasi-experimental technique used is shown in Table 2.

Table 2. Frequency of Quasi-Experiment Techniques

Quasi Experiment Technique	Frequency	Percentage
Nonequivalent Control Group Design	143	72.22
Factorial Design	47	23.74
Time Series Design	8	4.04
Counterbalance Design	0	0.00

Discussion

Some researchers should have explicitly mentioned the method of the summative evaluation carried out. Some researchers need to describe the formative evaluation explicitly and the summative evaluation. They only describe the evaluation steps carried out. Few researchers also include evaluation methods and techniques only at the end of development without explaining whether it is formative or summative. However, from the procedures and instruments used, the observer made inferences to identify the evaluation methods and techniques used, both as formative and summative evaluations. Several studies apply pre-experiment at the end of development and call it a summative evaluation. This kind of research is not analyzed because pre-experimental research is not included in the category of summative evaluation. This finding is not surprising because they are novice researchers with limited research experience. However, their evaluation procedure can be used as study material to formulate steps for coaching prospective teacher students in developing ICT-based learning media, especially in the summative evaluation section. On the other hand, many researchers have included clear and systematic evaluation methods and techniques, both formative and summative evaluation. Student variability in the potential, effort and environmental support is very influential. For diligent students, the opportunity for independent study from various sources on the internet, the opportunity to be involved in lecturer research, and the opportunity to participate in research competitions that are currently wide open have made them relatively mature in the field of research.

Quasi-experiment is the most chosen by researchers to conduct the summative evaluation. Researchers sometimes encounter difficulties in conducting randomization (Gay & Mills, 2012), both for ethical and other reasons (Indrayan, 2019), thus choosing a quasi-experimental, not a true experiment. Classes in schools have been formed in such a way according to applicable rules. Therefore, it is unethical for researchers to restructure the class on the grounds of randomizing students for research purposes. Another reason stated in the report is that they want to compare the performance or effectiveness of the developed ICT-based learning media. This decision is acceptable because the quasi-experiment is appropriate for comparisons (Plomp & Nieveen, 2013) and testing causal effects (nieveen & folmer, 2013).

There are several designs in the quasi-experimental, and it turns out that the non-equivalent control group design is the most chosen. Students, as researchers, did not explicitly mention why they chose the design. The observer analyzed student research reports and found that the technique was chosen for classic reasons, namely that many practical examples could be followed. Many studies that use non-equivalent control group designs make students familiar with the design, where familiarity tends to be the reason for choosing (khodabandelou & samah, 2012). Social learning theory, social behavior is learned by observing and imitating the behavior of others (bandura, 1977). Therefore, designs that researchers widely use tend to be followed by students. The commonly used models tend to be chosen by novice researchers (basu, 2018).

Factorial design and time series design are in second and third place, respectively, of the designs used by students in the summative evaluation of ICT-based learning media. Factorial design is widely used in experimental research. However, students who research summative evaluation of ICT-based learning media rarely involve other factors or variables, so they do not demand the application of factorial design. Time series design is appropriate for summative evaluation because it opens up opportunities for periodic observations. However, it must be admitted that not many studies have used it. As a result, not many examples can be used as guidelines by students. Even the counterbalance design has not yet been chosen in the summative evaluation of ICT-based learning media. Besides being considered complex, many relevant examples cannot be followed. Time, budget, and student skills are potential considerations. Surveys are also widely chosen in summative evaluation, either alone or as a support for quasi-experiments. As a supporter of quasi-experiments, surveys are used to obtain qualitative information on ICT-based learning media, both usability, practicality, and acceptability. The survey is also used separately if it is impossible to conduct a quasi-experiment due to time constraints. The survey can be used to obtain information that describes the characteristics of the sample in a relatively fast time (Ponto, 2015). Surveys can also be used to explore attitudes (Glasow, 2005). So that the survey approach can be applied to ICT-based learning media products that focus on learning in the attitude domain.

Surveys are mostly conducted with self-developed questionnaires and those developed by other researchers. One of the widely used questionnaires is the user experience questionnaire (UEQ) (Laugwitz et al., 2008) with 26 items to measure attractiveness, perspicuity, dependability, stimulation, and novelty. Besides UEQ, the instrument widely used in the summative evaluation of ICT-based learning media is learning object evaluation instruments (LORI). LORI measures content quality, learning goal alignment, feedback, adaptation, motivation, Ability to motivate and interest an identified population of learners, presentation design, interaction usability, accessibility, reusability, and standards compliance (Leacock & Nesbit, 2007). In addition, another instrument often used in learning media assessment surveys called The System Usability Scale (SUS), which consists of 10 items (Brooke, 2013). A case study is another approach used in the summative evaluation of ICT-based learning media, in addition to quasi-experiments and surveys. It is recognized that certain ICT-based learning media products are difficult to find test subjects to conduct quasi-experiments. ICT-based learning media products for children with special needs, for example, are quite difficult to get in an experimental class and a control class because the subjects are limited. Case studies are applied to ICT-based learning media products focusing on learning in the affective or psychomotor domain. In addition, ICT-based learning media products for elementary school children are also evaluated more summative with a case study approach in order to be able to observe in more detail the operational details of teaching materials and record the learning outcomes achieved (Rogaten et al., 2019). The affective and psychomotor domains be better evaluated using a longitudinal approach than comparison, which can be done through case studies (Symonds & Symonds, 2016).

4. CONCLUSION

Summative evaluation is carried out at the end of the development of ICT-based learning media to obtain information on whether to continue or stop development. The summative evaluation focuses on the effectiveness and usability of the media. Media is said to meet the effectiveness requirements if it can deliver students achieve the expected learning outcomes. Furthermore, usability requirements include functionality, acceptability, and ease of use. Functionality means that the media can carry out its functions properly, and acceptability means that users can accept the media. Ease of use means that the media is easy to operate for users. The intended users in this context are teachers or students. Summative evaluation can be done with a quasi-experimental approach, survey, or case study. Quasi-experiments are mainly used to evaluate effectiveness. If there are enough test subjects so that it is possible to form an experimental group and a control group of two cells each, the factorial method can be applied. If the test subjects are only allowed to form an experimental group and a control group for one cell each, the nonequivalent control group design and the counterbalanced design can be applied. The time series

method can be used if there is only one group. Quasi-experiments can be assisted by surveys or case studies to obtain a qualitative description of the effects of the intervention, such as functionality, acceptability, and ease of use.

5. ACKNOWLEDGMENT

First, we would like to thank the UNDIKSHA research institute that has funded this research. Furthermore, we would like to thank the UNDIKSHA library staff who have prepared the data for this research. Finally, we appreciate the contribution of the experts during the focus group discussion. Due to various limitations, this research can be continued by the next researcher, who is virtuous, so that more complete results are obtained.

6. REFERENCES

Aldoobie, N. (2015). ADDIE model. *American International Journal of Contemporary Research*, *5*(6), 68–72.

Allen, M. (2019). *Counterbalancing.* The SAGE Encyclopedia of Communication Research Methods.

Amin, Z., & Eng, K. H. (2003). Formative and Summative Assessment. *Basics in Medical Education*, 261–265. https://doi.org/10.1142/9789812795472_0028.

Bandura, A. (1977). *Social Learning Theory.* Prentice-Hall, Inc.

Basu, R. (2018). Instructional Design Models : Benefits and Challenges. *UGC Approved Journal*, *Vol.-41 N*(June), 31–36.

Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies*, *8*(2), 29–40.

Chang, S. L. (2006). The Systematic Design of Instruction. In *Educational Technology Research and Development* (Vol. 54, Issue 4). Harper Collins Publishers Inc. https://doi.org/10.1007/s11423-006-9606-0.

Creswell, J. W. (2015). A Concise Introduction to Mixed Methods Research. In *SAGE Publications, Inc.* (Vol. 1999, Issue December). Sage.

Deeley, S. J. (2018). Using technology to facilitate effective assessment for learning and feedback in higher education. *Assessment and Evaluation in Higher Education*, *43*(3), 439–448. https://doi.org/10.1080/02602938.2017.1356906.

Drljača, D., Latinović, B., Stanković, Ž., & Cvetković, D. (2017). ADDIE Model for Development of E-Courses. *In Documento Procedente de La International Scientific Conference on Information Technology and Data Related Research SINTEZA*, 242–247. https://doi.org/10.15308/sinteza-2017-242-247.

Earl, L. M. (2003). *Assessment As Learning: Using Classroom Assessment to Maximize Student Learning (Experts In Assessment Series)*. Corwin Press. http://www.amazon.com/Assessment-As-Learning-Classroom-Maximize/dp/0761946268.

Eraut, M. (1982). Evaluation a systematic approach. In *International Journal of Educational Development* (Vol. 2, Issue 3). SAGE. https://doi.org/10.1016/0738-0593(82)90010-4.

Farquhar, J., & Michels, N. (2016). *Triangulation without Tears, Marketing Challenges in a Turbulent Business Environment* (Issue 3). Heidelberg.

Fraenkel, Jack R., Wallen, N. E. (2009). How to Design and Evaluate Research in Education. In *McGraw-Hill Higher Education* (Issue 0). McGraw-Hill Inc.

Gall, M. D., Joyce P. Gall, & Borg., W. R. (2003). *Educational Research: An Introduction.* Pearson Education, Inc.

Gay, I. R., & Mills, E. G. (2012). Research method: Competencies for analysis and applications. In *Pearson Education, Inc* (Vol. 6, Issue 2). Pearson. http://library1.nida.ac.th/termpaper6/sd/2554/19755.pdf.

Glasow, P. A. (2005). *Fundamentals of survey research methodology*. MITRE, C3 Center. http://www.uky.edu/~kdbrad2/EPE619/Handouts/SurveyResearchReading.pdf.

Gravetter, F. J., Forzano, L. A. B. (2009). *Research Methods Behavioral Sciences*. WADSWORTH CENCAGE Learning.

Hattie, J. A., Brown, G. T. L., Ward, L., Irving, S. E., & Keegan, P. J. (2006). Formative Evaluation of an Educational Assessment Technology Innovation : Developers ' Insights into Assessment Tools for Teaching and Learning (asTTle). *Journal of Multidisciplinary Evaluation*, *5*(3), 1–54.

Heale, R., & Twycross., A. (2018). What is a case study? Roberta Heale. *Evid Based Nurs*, *21*.

Indrayan, A. (2019). Sampling and Sample Size. *Research Methods for Medical Graduates*, 147–168. https://doi.org/10.1201/9780429435034-8.

Joyce, A. (2019). *Formative vs. Summative Evaluation*. https://www.nngroup.com/articles/formative-vs-summative-evaluations/.

Khodabandelou, R., & Samah, S. A. A. (2012). Instructional Design Models for Online Instruction: From the Perspective of Iranian Higher Education. *Procedia - Social and Behavioral Sciences*, *67*(1), 545–552. https://doi.org/10.1016/j.sbspro.2012.11.359.

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a user experience questionnaire. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *5298 LNCS*, 63–76. https://doi.org/10.1007/978-3-540-89350-9_6.

Leacock, T. L., & Nesbit, J. C. (2007). A Framework for Evaluating the Quality of Multimedia Learning Resources. *Educational Technology & Society*, *10*(2), 44–59.

Loeb, S., Dynarski, S., Mcfarland, D., Morris, P., Reardon, S., & Reber, S. (2017). Descriptive Analysis in Education: A Guide for Researchers. NCEE 2017-4023. In *Eric*. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://eric.ed.gov/?id=ED573325.

LR, G., GE, M., & P, A. (2006). *Educational Research: Competencies for Analysis and Applications 8thEdition*. Pearson.

Mathiyazhagan, T., Nandan, D., & Khosla, M. (2010). Survey research method ABSTRACT. *Media Mimansa*, *4*(September), 34–82. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.464.5585&rep=rep1&type=pdf.

Mckenzie, B. K. (2006). Evaluating Educational Technologies : Technology Connoisseurs in the Campus *. *Annual Meeting of Eastern Educational Research Association*, *1998*, 1–10.

Miles, M. B., Huerman, A. M., & Johny Saldana. (2014). *Qualitatite Data Analysis: A Methods Sourcebook.* SAGE Publication, Inc.

Nassaji, H. (2015). Qualitative and descriptive research: Data type versus data analysis. *Language Teaching Research*, *19*(2), 129–132. https://doi.org/10.1177/1362168815572747.

Navarro-Rivera, J., & Kosmin, B. A. (2015). Surveys and Questionnaires. *The Routledge Handbook of Research Methods in the Study of Religion*. https://doi.org/10.4324/9780203154281.ch2_20.

Nieveen, N., & Folmer, E. (2013). Formative evaluation in educational design research. In *Educational Design Research. Part A: An introduction*. Institute for Curriculum Development (SLO). http://international.slo.nl/publications/edr/.

Patton, M. Q. (1994). The program evaluation standards: How to assess evaluations of educational programs. In *Evaluation Practice* (second, Vol. 15, Issue 2). SAGE. https://doi.org/10.1016/0886-1633(94)90009-4.

Plomp, T., & Nieveen, N. (2013). Educational Design Research Educational Design Research. In *Netherlands Institute for Curriculum Development: SLO*. Institute for Curriculum Development (SLO). http://www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ815766.

Ponto, PhD, APRN, AGCNS-BC, AOCNS®, J. (2015). Understanding and Evaluating Survey Research. *Journal of the Advanced Practitioner in Oncology*, *6*(2), 168. https://doi.org/10.6004/jadpro.2015.6.2.9.

Ponto. (2015). Understanding and Evaluating Survey Research. *Journal of the Advanced Practitioner in Oncology*, *6*(2), 168–171. https://doi.org/10.6004/jadpro.2015.6.2.9.

Popham, W. J. (1999). Modern educational measurement. In *Practical guidelines for the education leader. Michigan: Pearson* (Issue Pde 105). Prentice Hall, Inc.

Rogaten, J., Rienties, B., Sharpe, R., Cross, S., Whitelock, D., Lygo-Baker, S., & Littlejohn, A. (2019). Reviewing affective, behavioural and cognitive learning gains in higher education. *Assessment and Evaluation in Higher Education*, *44*(3), 321–337. https://doi.org/10.1080/02602938.2018.1504277.

Rowley, J. (2002). Using case studies in research. *Management Research News*, *25*(1), 16–27. https://doi.org/10.1108/01409170210782990.

Sarkies, M. N., Skinner, E. H., Bowles, K. A., Morris, M. E., Williams, C., O'Brien, L., Bardoel, A., Martin, J., Holland, A. E., Carey, L., White, J., & Haines, T. P. (2019). A novel counterbalanced implementation study design: Methodological description and application to implementation research. *Implementation Science*, *14*(1), 1–11. https://doi.org/10.1186/s13012-019-0896-0.

Scriven, M. (1992). Evaluation thesaurus. In *Choice Reviews Online* (Vol. 29, Issue 08). SAGE. https://doi.org/10.5860/choice.29-4272.

Spivey, G. (2007). A taxonomy for learning, teaching, and assessing digital logic design. In *Proceedings - Frontiers in Education Conference, FIE*. Addison Wesley Longman Inc. https://doi.org/10.1109/FIE.2007.4417846.

Springer, K. (2010). *Educational Research: a contextual approach.* Wiley.

Symonds, P., & Symonds, P. M. (2016a). The Case Study as a Research Method. *Case Studies*, 15–15. https://doi.org/10.4135/9781473915480.n2.

Symonds, P., & Symonds, P. M. (2016b). The Case Study as a Research Method. *Case Studies*, *9*, 15–15. https://doi.org/10.4135/9781473915480.n2.

Vedung, E. (2017). Public policy and program evaluation. In *Public Policy and Program Evaluation*. Transaction Publishers. https://doi.org/10.4324/9781315127767.

Verdinelli, S., & Scagnoli, N. I. (2013). Data display in qualitative research. *International Journal of Qualitative Methods*, *12*(1), 359–381. https://doi.org/10.1177/160940691301200117.

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting With Teacher Professional Development: Motives and Methods. *Educational Researcher*, *37*(8), 469–479. https://doi.org/10.3102/0013189X08327154.

Weston, T. (2004). Formative Evaluation for Implementation: Evaluating Educational Technology Applications and Lessons. *American Journal of Evaluation*, *25*(1), 51–64. https://doi.org/10.1177/109821400402500104.

Wüthrich, K. (2004). Triangulation in Social Research: Qualitative and Quantitative Methods Can Really Be Mixed. In *Developments in Sociology*. Causeway Press Ltd.

Zeelenberg, R., & Pecher, D. (2015). A method for simultaneously counterbalancing condition order and assignment of stimulus materials to conditions. *Behavior Research Methods*, *47*(1), 127–133. https://doi.org/10.3758/s13428-014-0476-9.