



## Formative Evaluation of Digital Learning Materials

Ni Made Sri Mertasari<sup>1\*</sup>, I Made Candiasa<sup>2</sup> 

<sup>1,2</sup> Program Studi Pendidikan Matematika, Universitas Pendidikan Ganesha, Singaraja, Indonesia

### ARTICLE INFO

#### Article history:

Received January 31, 2022

Revised February 11, 2022

Accepted July 12, 2022

Available online August 25, 2022

#### Kata Kunci:

Evaluasi, Formatif, Bahan Ajar, Digital

#### Keywords:

Evaluation, formative, learning materials, digital

#### DOI:

<https://doi.org/10.23887/jet.v6i3.44165>

### ABSTRAK

Penelitian ini bertujuan untuk merumuskan metode evaluasi bahan ajar digital yang relevan dengan model pengembangan yang digunakan. Upaya tersebut dilakukan karena ada indikasi bahwa mahasiswa calon guru masih sulit memilih metode evaluasi formatif dalam pengembangan bahan ajar digital. Upaya tersebut didahului dengan penelitian untuk mengidentifikasi model evaluasi bahan ajar digital yang sudah digunakan mahasiswa calon guru. Penelitian dilakukan secara kualitatif terhadap 488 dokumen laporan penelitian mahasiswa calon guru untuk lima tahun terakhir. Sekalipun terkadang tidak menyebut secara eksplisit, evaluasi formatif sudah dilakukan pada pengembangan bahan ajar digital, namun beberapa diantaranya belum dilandasi teori yang akurat. Metode evaluasi formatif yang dipilih beserta alasan pemilihannya dianalisis secara kualitatif lengkap dengan triangulasi dan hasilnya dibahas dalam focus group discussion. Sampel diambil dari hasil penelitian calon guru. Expert review dilakukan untuk uji validitas isi dan uji desain. Uji validitas isi dapat dilakukan secara kualitatif atau kuantitatif dengan menerapkan teknik Lawshe, Aiken, atau Gregory. Uji desain cukup dilakukan dengan uji black-box. Uji white-box sangat jarang yang menerapkan dan direkomendasikan untuk diterapkan bila pengembangan menggunakan bahasa pemrograman. Uji one-to-one dan small group dilakukan dengan teknik think aloud, cognitive walk through, atau heuristik. Hasil analisis menemukan bahwa metode evaluasi formatif bahan ajar yang diperlukan adalah expert review, one-to-one evaluation, small group evaluation, dan field test. Selanjutnya untuk field test dapat diterapkan metode inquiry dengan berbagai teknik, seperti, learning object review instrument (LORI), user experience questionnaire (UEQ), atau system usability scale (SUS).

### ABSTRACT

The use of digital learning materials through e-learning is expected to give students experience to be able to learn themselves in the industrial era 4.0. This study aims to formulate an evaluation method for digital learning materials that is relevant to the development model used. A qualitative approach with an evaluative design through in-depth analysis is applied to student research reports, specifically research on the development of digital learning materials. The formative evaluation method chosen and the reasons for the selection were analyzed and the results were discussed in a focus group discussion. The sample was taken from the results of research by prospective teacher. Expert review was conducted to test the validity of the content and test the design. Content validity test can be done qualitatively or quantitatively by applying Lawshe, Aiken, or Gregory techniques. The design test was carried out with a black-box test. White-box testing is very rarely implemented and is recommended when developing using a programming language. One-to-one and small group tests were conducted using think aloud, cognitive walk through, or heuristics techniques. The results of the analysis found that the formative evaluation methods of teaching materials needed were expert review, one-to-one, small group, and field evaluation. Furthermore, for field tests, inquiry methods can be applied with various techniques, such as learning object review instrument (LORI), user experience questionnaire (UEQ), or system usability scale (SUS).

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.  
Copyright © 2022 by Author. Published by Universitas Pendidikan Ganesha.



## 1. INTRODUCTION

The paradigm shift in classroom learning has occurred several years ago. Learning does not only utilize information sourced from books, but from various sources in the form of digital learning materials through e-learning. Through e-learning students learn comfortably by communicating with each other on-line (Avando Bastari et al., 2021; Carvalho et al., 2020; X. Wu et al., 2022). In addition, e-learning has a multi-layer network that supports deep and powerful learning through text mining (Kundu & Bej, 2021; Vagg et al., 2020; Young et al., 2018). E-learning opens opportunities for students to learn from various sources, so that they can learn according to their respective characteristics to obtain more optimal learning outcomes. Learning materials can be presented in various forms, such as text, images, graphics, videos, or animations that students can choose according to their wishes (Coman et al., 2020; M Nurtanto et al., 2020; Rodríguez-Gómez et al., 2014). Digital learning materials in e-learning provide opportunities for students to communicate learning without being bound by time and geographical distance by utilizing technology, especially the internet. E-learning provides

\* Corresponding Author: Ni Made Sri Mertasari: [srimerasari@undiksha.ac.id](mailto:srimerasari@undiksha.ac.id)

opportunities for learner control, where students can learn according to their experiences and control their own learning direction (Ferdianto & Dwiniasih, 2019; Öz, 2014; Yilmaz, 2017). In addition, e-learning also opens up opportunities for students to collaborate online to form an online learning community (Ouyang & Scharber, 2017; J. Y. Wu & Nian, 2021). Thus, the use of digital learning materials through e-learning is expected to give students experience to be able to learn themselves in the industrial era 4.0, which is marked, among others, by the use of the Internet of Things (Morrar et al., 2017; Roblek et al., 2016).

Many digital learning materials have been developed, either by companies engaged in education or educational institutions. Teachers and even student teacher candidates have also been heavily involved in the development of digital learning materials, both individually and in teams. Digital learning materials are developed in various forms, such as multimedia, hypermedia, web, video, animation, presentation, or just text documents (Dewi & Alam, 2020; Suryani, 2018; Suwartono & Aniuranti, 2019). The packaging of digital learning materials also varies, such as stand-alone applications, online applications, or materials on the website (Davis et al., 2018; Rajendran & Santhanam, 2021). There are many models for developing digital learning materials and each of them has various stages of development. In general, although with relatively different terms, there are three stages of development, namely analysis or preliminary, prototype, and evaluation or assessment (Kljun et al., 2020; Tjeerd Plomp, 2013).

One of the problems encountered by prospective teacher students or even teachers in the development of digital learning materials is in formative evaluation. It is recognized that the evaluation of digital learning materials is very complex and time-consuming (Adzobu, N, 2014; Hughes et al., 2020). Evaluation in e-learning programs is the weakest component, one of the suspected causes is the absence of standards. In fact, on the other hand, there are many computer application evaluation models, both theoretically and in best practice. In addition to adopting an evaluation model of learning media in general, there is an opportunity to adopt an evaluation process on software engineering into a formative evaluation of digital learning materials (Muhammad Nurtanto et al., 2020; Oktiningrum & Wardhani, 2020; Schneider & Bodensohn, 2017).

Evaluation is a systematic process for planning and studying the impact of a policy, performance, program, or initiative to create, promote, or sustain social change (Abrahams & Reiss, 2012; Okoye, 2014). Evaluation in the development of digital learning materials is the process of reviewing the developed learning materials and the results obtained from the learning process (Hoang & Arch-Int, 2013; Martin et al., 2019). The development of digital learning materials is said to be successful if the learning has achieved the desired results. Many types of evaluation are well known, among which the most popular is formative evaluation. Formative evaluation refers to a series of evaluations carried out by teachers during the learning process in order to modify learning programs to improve student achievement (Hughes et al., 2020; Maier et al., 2016; Wongwatkit et al., 2016). Formative evaluation is carried out during the learning process to get feedback from students which is used as information to make improvements to the learning program. Previous study illustrates that formative assessment is when the cook tastes the soup, while summative assessment is when the customer tastes the soup (Madaus & Kellaghan, 2000). Formative evaluation can be divided into three sequential steps, namely one-to-one evaluation, small group evaluation, and field trial evaluation (Bhat & Bhat, 2019; Github, 2013; Tessmer, 2013).

In accordance with the development of science and technology, the term formative evaluation is also used in the development of instructional designs. It is in line with previous study that state formative evaluation in the development of learning materials is interpreted as collecting data or information during the process of developing learning products that are used to increase learning effectiveness (Aldoobie, 2015; Tjeerd Plomp, 2013). Moreover other previous research also emphasizes that the function of formative evaluation is 'to improve', in this case improving the quality of learning materials (Schneider & Bodensohn, 2017). Therefore, formative evaluation focuses on uncovering deficiencies that exist during the development process with the aim of generating suggestions to improve them. In the end, the product of learning materials developed can meet the needs of users, which is effective in achieving learning outcomes and efficient in the use of time and learning facilities. This study tries to analyze the formative evaluation of digital learning materials developed by prospective teachers, in an effort to obtain an adequate formulation of the formative evaluation model of digital learning materials.

## 2. METHOD

This study uses a qualitative approach with an evaluative design through in-depth analysis of a case (Creswell, 2014; Morgan, 2022). The research population is a research report document on the development of digital teaching materials produced by prospective teacher students. The sample was taken from the results of research by prospective teacher students at Ganesha University of Education for the last five years. Data was collected through observation of research report documents to identify and describe the evaluation technique of digital learning materials developed. Triangulation in data collection is done by involving two observers to

obtain data from different perspectives. The two observers worked independently, with occasional discussions. Such triangulation can help ensure that underlying biases arising from the use of single observers are addressed (Farquhar & Michels, 2016; Noble & Heale, 2019). Thus the theory that was successfully built from the collected data became stronger.

The data obtained were analyzed qualitatively using technique which consists of three steps, namely data condensation, data display, and drawing and verifying conclusions (Miles et al., 2014; Susilawati et al., 2019). During data condensation, data is selected which is a keyword so that the data is simpler and more focused on the research objectives. Furthermore, when displaying data, organizing data is carried out so that the information contained in it is properly arranged, making it easier to draw conclusions. Visual displays are very helpful in presenting data for inference (Istri Aryani & Rahayuni, 2016; Verdinelli & Scagnoli, 2013). The data is organized in the form of a matrix, in order to show the relationship between the relevant information. After looking at the series of information in the matrix, conclusions are then drawn.

Conclusions are drawn by reasoning from the data that has been organized. If the initial conclusions obtained are still considered weak, then re-access to data storage is carried out for further analysis. Re-analysis can be from data condensation or simply data that has been organized in the data display. If the existing data is deemed insufficient, then the data can be retrieved again. To strengthen the conclusions obtained, verification is carried out through theoretical triangulation, by interpreting using the perspectives of several theories (Farquhar & Michels, 2016; Lubis & Dasopang, 2021).

### 3. RESULT AND DISCUSSION

#### Result

Observations on research reports on the development of digital learning materials produced by prospective student teachers in university library information systems found that the existence of formative evaluation varied widely. There are research reports that describe formative evaluation in full and in detail, there are research reports that only mention formative evaluation in outline, and there are even reports that do not include formative evaluation. The last group of research reports were not included in further observations because they did not contain the necessary information. Observations looked at the formative evaluation methods and techniques used and obtained information as listed in the Table 1.

**Table 1.** Frequency of Formative Evaluation Method

Formative Evaluation Method	Frequency
One-to-one, small group, and field trial	78 (16%)
One-to-one and small group	49 (10%)
One-to-one and field trial	39 (8%)
Small group and field trial	83 (17%)
Only one-to-one	49 (10%)
Only small group	88 (18%)
Only small group	102 (21%)

Based on Table 1, the evaluation techniques applied to each formative evaluation method vary widely. Some reports do not contain the evaluation technique used, but contain a description. Based on the description included, the observer classifies the technique used according to the relevant theory. The results of the analysis found that the one-to-one evaluation was dominated by the use of think aloud techniques, both concurrent think aloud and retrospective think aloud. The think aloud technique was used in 99.2% of products, while the remaining 0.8% used the first click technique. The first click technique is used in digital learning material products developed for students with special needs. Furthermore, for small group evaluation, 97.6% of products were evaluated using cognitive-walkthrough techniques, and the remaining 2.4% using heuristic techniques. Some products even apply both. For the evaluation of the field trial, all products that evaluate the field trial apply the inquiry technique. The inquiry instruments used varied, including User Experience Questionnaire, System Usability Scale, Learning Object Review Instrument, and a questionnaire (Brooke, 2013; Vargo et al., 2003). Some of the products were even evaluated with the inquiry technique using an instrument developed by the researcher himself.

One interesting thing found in several digital learning products is the evaluation of the validity of the material in 18% of the total observed products. Evaluations are carried out by relevant experts with qualitative, quantitative approaches or even those who do both. In the qualitative approach, the expert makes corrections to the material followed by providing revisions to the necessary sections or providing development recommendations if needed. In the quantitative approach, an expert group consisting of at least two experts

provides an assessment of the material using a prepared worksheet. Data analysis used a variety of techniques, including Content Validity, the Item-Goal Congruence Index formula, the Aiken formula or the Gregory formula (Candra & Retnawati, 2020; García-Ceberino et al., 2020). Material validity testing is generally carried out at the end of the development phase. As a result, if a material error occurs, both the analysis phase and the development phase must be repeated.

Another finding in the formative assessment is the existence of functional tests and structural tests of learning application products. Although the percentage is very small, which is under 5% for functional tests and under 1% for structural tests of the observed products, these findings are important to be observed in the evaluation of learning applications in the future. Functional testing and structural testing are stages of software engineering in terms of testing the accuracy of software functions and structures. Functional tests are carried out to see whether the developed learning material products have functioned as specified, while structural tests are carried out to see the correctness of the structure of the developed learning material products. It seems that the functional tests and structural tests that have been carried out have not fully met the demands of software engineering, but at least efforts in that direction have begun. Functionality test is done by black-box testing, while structural test is done by white-box testing. Black-box testing and white-box testing are not alternatives, but complement each other to reveal undetected errors in their respective tasks.

## Discussion

Think aloud is the testing technique chosen by the majority of researchers for the one-to-one test. Think Aloud is done by asking users to use the application while continuing to think aloud and express their thoughts as they move through the interface. While operating the application the user verbalizes continuously about what they think when using the application. Verbalization made by the user is used as feedback by the observer to interpret the problems that occur in the application. All activities when the user verbalizes will be recorded, so that the actual events experienced by the user can be known (Lyon et al., 2021; Weninger et al., 2020). Thus, the purpose of a one-to-one evaluation, which is to seek clarity, impact, and feasibility, can be met (Quinõnes & Rusu, 2017; Vieira et al., 2019; Yeni & Cagiltay., 2017). Think aloud technique divided into two, namely concurrent think aloud (CTA) and retrospective think aloud (RTA). In CTA, users are asked to verbalize their thoughts during the execution of the given scenario tasks, while in RTA users are asked to verbalize after carrying out the given scenario tasks. CTA has the opportunity to get more problems through observation because users are monitored directly while using the application. On the other hand, the RTA has a greater chance of having problems detected by verbalization.

The application of CTA provides several advantages, namely being able to understand directly how when users are working on scenarios, the evaluator can find out directly about what the user feels and how the user thinks when working on these scenarios. The drawback of CTA is that it takes a very long time, so it can make users feel tired of the tests carried out. On the other hand, RTA also has the advantage that it allows users to focus on working on scenarios. Another advantage of RTA is that it is more efficient in testing time. There is a time lag between implementation and verbalization which can result in a change in the user's mind, so the explanation given can be different. These factors can be a drawback of RTA. In addition to the think aloud technique, one-to-one evaluation can also be done using the first click technique. First click testing evaluates the user's behavior the first time using the application. One of the most significant advances in usability testing was the focus on the user's first click in each scenario, especially when participants interacted with websites (Dudek & Heiser, 2017; Vargo et al., 2003). The success of users using the application is closely related to the first time they use the application. If the user's first click is correct, then the probability of getting the overall scenario correct is 0.87, whereas if the first click is incorrect, then the probability of finally getting the scenario correct is only 0.46.

Although no researchers have used it yet, the eye tracking technique can also be applied to a one-to-one test. Eye tracking considers the application of eye movements to user interfaces, both for analyzing the interface, measuring usability, and gaining insight into human performance, as well as as a medium of actual control in human-computer dialogue (Bojko, 2013; Jacob & Karn, 2003). The eye tracking method is used to find out whether the respondent has difficulty opening a web page, knowing how much effort is needed when searching on a web page by reading the direction of the respondent's eyes when viewing objects on the monitor screen. Following the respondent's eyes at work makes the researcher feel as if he is in the respondent's head and thinking with him, so that it can be known in more detail which parts can be used and which parts cannot be used. Faulty or confusing parts of the web or system interface can be identified, so that corrective action can be taken.

It is in line with previous study that the purpose to clarify the terminology related to formative assessment and its usage (Srivastava et al., 2018). The result found that the use of formative assessments, or other diagnostic efforts within classrooms, provides information that should help facilitate improved pedagogical practices and instructional outcomes. However, a review of the formative assessment literature revealed that



there is no agreed upon lexicon with regard to formative assessment and suspect methodological approaches in the efforts to demonstrate positive effects that could be attributed to formative assessments. Moreover it is also reinforced by other study that investigates the effectiveness of using the classroom response system "Plickers," which is a technology-based formative assessment tool, in improving students' learning (Elmahdi et al., 2018). The findings revealed that students believe in the importance of formative assessment and receiving immediate feedback which is supported by the use of "Plickers." Moreover, the findings showed that using "Plickers" for formative assessment aids the learning process as it improves students' participation, saves the learning time, guarantees equal participation opportunities, and creates a fun and exciting learning environment. Formative evaluation is expected to be able to provide feedback carefully and quickly, so that the digital teaching material products obtained are more effective and efficient. The implication of this research is to provide information related to the formative evaluation of digital learning materials developed by prospective teachers, and to obtain an adequate formulation of the formative evaluation model of digital learning materials. This will be very useful for educators in applying formative assessment in learning. This research is far from perfect, especially in the area of participation coverage which still takes samples from only one agency. It is hoped that the next research will be able to expand the scope and deepen the discussion related to the formative evaluation of digital learning materials.

#### 4. CONCLUSION

Formative evaluations have been carried out in the development of digital teaching materials by prospective teacher students, although improvement efforts are still needed. In accordance with the provisions, formative evaluations have been carried out at each stage of development. In general, there are three stages of development, namely analysis or preliminary, prototype, and evaluation or assessment. The preliminary stage requires an evaluation of content validity, both qualitatively and quantitatively. Furthermore, the prototype stage requires a one-to-one evaluation, small group, and field trial. Evaluation generally uses the think aloud first click technique. Furthermore, small group evaluation can apply cognitive-walkthrough and heuristic techniques. The most common field trial evaluation is the inquiry technique. The inquiry instruments used varied, including the User Experience Questionnaire (UEQ), System Usability Scale (SUS), Learning Object Review Instrument (LORI) and questionnaires from various sources or those developed by the researchers themselves. Many formative evaluation techniques were adopted from software evaluation techniques in general. This makes sense because the development of digital learning materials is a special case of software engineering.

#### 5. REFERENCES

- Abrahams, I., & Reiss, M. J. (2012). Practical work: Its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching*, 49(8). <https://doi.org/10.1002/tea.21036>.
- Adzobu, N. Y. (2014). Design, Use and Evaluation of E-Learning Platforms: Experiences and Perspectives of a Practitioner from the Developing World Studying in the Developed World. *Informatics*, 1, 147–159. <https://doi.org/10.3390/informatics1020147>.
- Aldoobie, N. (2015). ADDIE Model. *American International Journal of Contemporary Research*, 5(6). [www.aijcrnet.com/journals/Vol\\_5\\_No\\_6\\_December\\_2015/10.pdf](http://www.aijcrnet.com/journals/Vol_5_No_6_December_2015/10.pdf).
- Avando Bastari, Adi Bandono, & Okol Sri Suharyo. (2021). The development strategy of smart campus for improving excellent navy human resources. *Global Journal of Engineering and Technology Advances*, 6(2), 033–043. <https://doi.org/10.30574/gjeta.2021.6.2.0011>.
- Bhat, B. A., & Bhat, G. J. (2019). Formative and summative evaluation techniques for improvement of learning process. *European Journal of Business & Social Sciences*, 7(5), 776–785. [https://www.researchgate.net/profile/Bilal-Bhat-6/publication/333633265\\_Formative\\_and\\_Summative\\_Evaluation\\_Techniques\\_for\\_Improvement\\_of\\_Learning\\_Process/links/5e6b7e48a6fdccf321d9926e/Formative-and-Summative-Evaluation-Techniques-for-Improvement-of-Learning-Process.pdf](https://www.researchgate.net/profile/Bilal-Bhat-6/publication/333633265_Formative_and_Summative_Evaluation_Techniques_for_Improvement_of_Learning_Process/links/5e6b7e48a6fdccf321d9926e/Formative-and-Summative-Evaluation-Techniques-for-Improvement-of-Learning-Process.pdf).
- Bojko, A. (2013). *Eye Tracking The User Experience*. Rosenfeld Media.
- Brooke, J. (2013). SUS: a retrospective. *Journal of Usability Studies*, 8(2), 29–40. [https://www.researchgate.net/profile/John-Brooke-6/publication/285811057\\_SUS\\_a\\_retrospective/links/5ee5c4a792851ce9e7e38a75/SUS-a-retrospective.pdf](https://www.researchgate.net/profile/John-Brooke-6/publication/285811057_SUS_a_retrospective/links/5ee5c4a792851ce9e7e38a75/SUS-a-retrospective.pdf).
- Candra, & Retnawati, H. (2020). A meta-analysis of constructivism learning implementation towards the learning outcomes on civic education lesson. *International Journal of Instruction*, 13(2), 835–846. <https://doi.org/10.29333/iji.2020.13256a>.

- Carvalho, J. V., Pereira, R. H., & Rocha, Á. (2020). A systematic literature review on maturity models for information systems in higher education institutions. *Innovations in Education and Teaching International*, 57(4), 434–449. <https://doi.org/10.1080/14703297.2019.1648219>.
- Coman, C., Țiru, L. G., Meseșan-Schmitz, L., Stanciu, C., & Bularca, M. C. (2020). Online teaching and learning in higher education during the coronavirus pandemic: Students' perspective. *Sustainability (Switzerland)*, 12(24), 1–22. <https://doi.org/10.3390/su122410367>.
- Creswell, J. W. (2014). *E Book Research Design Cressweell 2014* (4th Editio). SAGE Publications, Inc.
- Davis, D., Chen, G., Hauff, C., & Houben, G. J. (2018). Activating learning at scale: A review of innovations in online learning strategies. *Computers and Education*, 125(June), 327–344. <https://doi.org/10.1016/j.compedu.2018.05.019>.
- Dewi, E. R., & Alam, A. A. (2020). Transformation Model for Character Education of Students. *Cypriot Journal of Educational Sciences*, 15(5), 1228–1237. <https://doi.org/10.18844/CJES.V15I5.5155>.
- Dudek, J., & Heiser, R. (2017). Elements, principles, and critical inquiry for identity-centered design of online environments. *Journal of Distance Education*, 32(2), 1–18. <https://search.proquest.com/openview/36da03443bbbdec42efcb9d7d1ceb2d8/1?pq-origsite=gscholar&cbl=446313>.
- Elmahdi, I., Al-Hattami, A., & Fawzi, H. (2018). Using Technology for Formative Assessment to Improve Students' Learning. *Turkish Online Journal of Educational Technology-TOJET*, 17(2), 182–188. <https://eric.ed.gov/?id=EJ1176157>.
- Farquhar, J., & Michels, N. (2016). Triangulation without tears. *In Marketing Challenges in a Turbulent Business Environment*, 325–330. [https://doi.org/10.1007/978-3-319-19428-8\\_86](https://doi.org/10.1007/978-3-319-19428-8_86).
- Ferdianto, F., & Dwiniasih. (2019). Learning Management System (LMS) schoology: Why it's important and what it looks like. *Journal of Physics: Conference Series*, 1360(1). <https://doi.org/10.1088/1742-6596/1360/1/012034>.
- García-Ceberino, J. M., Antúnez, A., Ibáñez, S. J., & Feu, S. (2020). Design and Validation of the Instrument for the Measurement of Learning and Performance in Football. *International Journal of Environmental Research and Public Health*, 17(13), 4629. <https://doi.org/10.4135/9781506326139.n240>.
- Githua, B. N. (2013). Secondary school students' perceptions of mathematics formative evaluation and the perceptions' relationship to their motivation to learn the subject by gender in Nairobi and Rift Valley Provinces, Kenya. *Asian Journal of Social Sciences and Humanities*, 2(1), 174–183. [http://www.ajssh.leena-luna.co.jp/AJSSHPDFs/Vol.2\(1\)/AJSSH2013\(2.1-19\).pdf](http://www.ajssh.leena-luna.co.jp/AJSSHPDFs/Vol.2(1)/AJSSH2013(2.1-19).pdf).
- Hoang, L. P., & Arch-Int, N. (2013). Assessment of open-ended questions using a multidimensional approach for the interaction and collaboration of learners in e-learning environments. *Journal of Universal Computer Science*, 19(7), 932–949. [https://www.researchgate.net/profile/Loc-Hoang-Phuoc/publication/286190008\\_Assessment\\_of\\_Open-Ended\\_Questions\\_using\\_a\\_Multidimensional\\_Approach\\_for\\_the\\_Interaction\\_and\\_Collaboration\\_of\\_Learners\\_in\\_E-Learning\\_Environments/links/584e066808aed95c25032e12/Ass](https://www.researchgate.net/profile/Loc-Hoang-Phuoc/publication/286190008_Assessment_of_Open-Ended_Questions_using_a_Multidimensional_Approach_for_the_Interaction_and_Collaboration_of_Learners_in_E-Learning_Environments/links/584e066808aed95c25032e12/Ass).
- Hughes, M., Salamonson, Y., & Metcalfe, L. (2020). Student engagement using multiple-attempt 'Weekly Participation Task' quizzes with undergraduate nursing students. *Nurse Education in Practice*, 46(August 2018), 102803. <https://doi.org/10.1016/j.nepr.2020.102803>.
- Istri Aryani, I. G. A., & Rahayuni, N. K. S. (2016). Innovation of Teaching and Learning English Applied to Animal Sciences' Student with the Combination of Computer Media and Audio Visual. *International Journal of Linguistics, Literature and Culture*, 2(1). <https://doi.org/10.21744/ijllc.v2i1.5>.
- Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *In The Mind's Eye*, 573–605. <https://doi.org/10.1016/B978-044451020-4/50031-1>.
- Kljun, M., Geroimenko, V., & Čopič Pucihar, K. (2020). Augmented Reality in Education: Current Status and Advancement of the Field. *In Springer Series on Cultural Computing*. [https://doi.org/10.1007/978-3-030-42156-4\\_1](https://doi.org/10.1007/978-3-030-42156-4_1).
- Kundu, A., & Bej, T. (2021). Experiencing e-assessment during COVID-19: an analysis of Indian students' perception. *Higher Education Evaluation and Development*, 15(2), 114–134. <https://doi.org/10.1108/heed-03-2021-0032>.
- Lubis, A. H., & Dasopang, M. D. (2021). Online Learning During the Covid-19 Pandemic: How is it Implemented in Elementary Schools? *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, 11(1), 120. <https://doi.org/10.25273/pe.v11i1.8618>.
- Lyon, A. R., Coifman, J., Cook, H., McRee, E., Liu, F. F., Ludwig, K., & McCauley, E. (2021). The Cognitive Walkthrough for Implementation Strategies (CWIS): a pragmatic method for assessing implementation strategy usability. *Implementation Science Communications*, 2(1), 1–16. <https://doi.org/10.1186/s43058-021-00183-0>.
- Madaus, G. F., & Kellaghan, T. (2000). Models, metaphors, and definitions in evaluation. *In Evaluation Models*,

- 19–31. [https://link.springer.com/content/pdf/10.1007/0-306-47559-6\\_2.pdf](https://link.springer.com/content/pdf/10.1007/0-306-47559-6_2.pdf).
- Maier, U., Wolf, N., & Randler, C. (2016). Effects of a computer-assisted formative assessment intervention based on multiple-tier diagnostic items and different feedback types. *Computers and Education*, 95, 85–98. <https://doi.org/10.1016/j.compedu.2015.12.002>.
- Martin, F., Ritzhaupt, A., Kumar, S., & Budhrani, K. (2019). Award-winning faculty online teaching practices: Course design, assessment and evaluation, and facilitation. *Internet and Higher Education*, 42(March), 34–43. <https://doi.org/10.1016/j.iheduc.2019.04.001>.
- Miles, B., Huberman, M., & Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook*. SAGE Pub.
- Morgan, H. (2022). Conducting a Qualitative Document Analysis. *The Qualitative Report*, 27(1), 64–77. <https://doi.org/10.46743/2160-3715/2022.5044>.
- Morrar, R., Arman, H., & Mousa, S. (2017). The fourth industrial revolution (Industry 4.0): A social innovation perspective. *Technology Innovation Management Review*, 7(11), 12–20. [https://timreview.ca/sites/default/files/Issue\\_PDF/TIMReview\\_November2017.pdf#page=12](https://timreview.ca/sites/default/files/Issue_PDF/TIMReview_November2017.pdf#page=12).
- Noble, H., & Heale. (2019). Triangulation in research, with examples. *Evidence-Based Nursing*, 22(3). <https://doi.org/10.1136/ebnurs-2019-103145>
- Nurtanto, M, Sofyan, H., & Pardjono, P. (2020). E-Learning Based AutoCAD 3D Interactive Multimedia on Vocational Education (VE) Learning. *Journal of Engineering Education*. <https://doi.org/10.16920/jeet/2021/v34i4/155014>.
- Nurtanto, Muhammad, Pardjono, P., Widarto, W., & Ramdani, S. D. (2020). The effect of STEM-EDP in professional learning on automotive engineering competence in vocational high school. *Journal for the Education of Gifted Young Scientists*, 8(2), 633–649. <https://doi.org/10.17478/JEGYS.645047>.
- Okoye, M. D. (2014). Authentic Assessment and Evaluation: Paramount Means for the Maximization of Teaching and Learning. *Journal of Educational and Social Research*, 4(7), 31–40. <https://doi.org/10.5901/jesr.2014.v4n7p31>.
- Oktiningrum, W., & Wardhani, D. A. P. (2020). Developing HOT's mathematics task with Indonesian heritage as context to assess mathematical literacy of students in primary school. *International Journal for Educational and Vocational Studies*, 1(8), 69. <https://doi.org/10.29103/ijevs.v2i1.1997>.
- Ouyang, F., & Scharber, C. (2017). The influences of an experienced instructor's discussion design and facilitation on an online learning community development: A social network analysis study. *Internet High. Educ.*, 35, 34–47. <https://doi.org/10.1016/j.iheduc.2017.07.002>.
- Öz, H. (2014). Pre-service English Teachers' Perceptions of Web-based Assessment in a Pedagogical Content Knowledge Course. *Procedia - Social and Behavioral Sciences*, 141, 45–58. <https://doi.org/https://doi.org/10.1016/j.sbspro.2014.05.010>.
- Quinões, D., & Rusu, C. (2017). How to develop usability heuristics: A systematic literature review. *Computer Standards & Interfaces*, V., 53, 89–122. <https://doi.org/10.1016/j.csi.2017.03.009>.
- Rajendran, D., & Santhanam, P. (2021). Towards digital game-based learning content with multi-objective reinforcement learning. *Materials Today: Proceedings*, xxx. <https://doi.org/10.1016/j.matpr.2021.03.156>.
- Roblek, V., Meško, M., & Krapež, L. (2016). "A Complex View of Industry 4.0." *SAGE Open*, 6(2), 2158244016653987. <https://doi.org/10.1177/2158244016653987>. [sgo.sagepub.com](http://sgo.sagepub.com).
- Rodríguez-Gómez, G., Quesada-Serra, V., & Ibarra-Sáiz, M. S. (2014). Learning-Oriented E-Assessment: the Effects of A Training and Guidance Programme on Lecturers' Perceptions. *Assessment and Evaluation in Higher Education*, 1–18. <https://doi.org/http://dx.doi.org/10.1080/02602938.2014.979132>.
- Schneider, C., & Bodensohn, R. (2017). Student Teachers' Appraisal of the Importance of Assessment in Teacher Education and Self-Reports on the Development of Assessment Competence. *Assessment in Education: Principles, Policy and Practice*, 24(2), 127–146. <https://doi.org/http://dx.doi.org/10.1080/0969594X.2017.1293002>.
- Srivastava, T. K., Mishra, V., & Waghmare, L. S. (2018). Formative Assessment Classroom Techniques (FACTs) for better learning in pre-clinical medical education: A controlled trial. *Journal of Clinical and Diagnostic Research*, 12(9). [https://www.researchgate.net/profile/Tripti-Srivastava/publication/327229061\\_Formative\\_Assessment\\_Classroom\\_Techniques\\_FACTs\\_for\\_Better\\_Learning\\_in\\_Pre-Clinical\\_Medical\\_Education\\_A\\_Controlled\\_Trial/links/5b8673f7299bf1d5a72eef8e/Formative-Assessment-Classroom-Techniques-FACTs-for-Better-Learning-in-Pre-Clinical-Medical-Education-A-Controlled-Trial.pdf](https://www.researchgate.net/profile/Tripti-Srivastava/publication/327229061_Formative_Assessment_Classroom_Techniques_FACTs_for_Better_Learning_in_Pre-Clinical_Medical_Education_A_Controlled_Trial/links/5b8673f7299bf1d5a72eef8e/Formative-Assessment-Classroom-Techniques-FACTs-for-Better-Learning-in-Pre-Clinical-Medical-Education-A-Controlled-Trial.pdf).
- Suryani, N. (2018). Utilization of digital media to improve the quality and attractiveness of the teaching of history. *The 2nd International Conference On Teacher Training and Education*, 2(1), 131–144. <https://jurnal.uns.ac.id/iccte/article/view/8164>.
- Susilawati, W. O., Widodo, H., & Sumarno, S. (2019). Strategy of Teachers in Supporting Environmentally

- Sustainable Development. *Journal of Education and Learning (EduLearn)*, 13(2), 247–254. <https://doi.org/https://doi.org/10.11591/edulearn.v13i2.12167>.
- Suwartono, T., & Aniranti, A. (2019). Digital Teaching Tools in 21st Century EFL Classroom: Are Our Teachers Ready? *ELLITE: Journal of English Language, Literature, and Teaching*, 3(2), 57. <https://doi.org/10.32528/ellite.v3i2.1916>.
- Tessmer, M. (2013). *Planning and conducting formative evaluations*. Routledge.
- Tjeerd Plomp. (2013). Educational Design Research: an Introduction. In T. Plomp & N. Nieveen (Eds.), *An Introduction to Educational Design Research* (pp. 9–36). Netzdruk, Enschede. <https://www.academia.edu/download/48472657/educational-design-research-part-a.pdf#page=12>.
- Vagg, T., Balta, J. Y., Bolger, A., & Lone, M. (2020). Multimedia in Education: What do the Students Think? *Health Professions Education*, 6(3), 325–333. <https://doi.org/doi.org/10.1016/j.hpe.2020.04.011>.
- Vargo, J., Nesbit, J. C., Belfer, K., & Archambault, A. (2003). Learning object evaluation: Computer mediated collaboration and inter-rater reliability. *International Journal of Computers and Applications*, 25(3), 1–8. <https://doi.org/10.1080/1206212X.2003.11441703>.
- 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>.
- Vieira, E. A. O., Silveira, A. C. D., & Martins, R. X. (2019). Heuristic evaluation on usability of educational games: A systematic review. *Informatics in Education*, 18(2), 427–442. <https://doi.org/10.15388/infedu.2019.20>.
- Weninger, M., Grünbacher, P., Gander, E., & Schörgenhumer, A. (2020). Evaluating an interactive memory analysis tool: Findings from a cognitive walkthrough and a user study. *Proceedings of the ACM on Human-Computer Interaction*, 1–37. <https://doi.org/10.1145/3394977>.
- Wongwatkit, C., Srisawasdi, N., Hwang, G. J., & Panjaburee, P. (2016). Influence of an Integrated Learning Diagnosis and Formative Assessment-based Personalized We Learning Approach on Students Learning Performances and Perceptions. *Interactive Learning Environments*, 1–15. <https://doi.org/http://dx.doi.org/10.1080/10494820.2016.1224255>.
- Wu, J. Y., & Nian, M. W. (2021). The dynamics of an online learning community in a hybrid statistics classroom over time: Implications for the question-oriented problem-solving course design with the social network analysis approach. *Computers & Education*, 166, 104120. <https://doi.org/10.1016/j.compedu.2020.104120>.
- Wu, X. ., He, Z. ., Li, M. ., Han, Z. ., & Huang, C. (2022). Identifying Learners' Interaction Patterns in an Online Learning Community. *Int. J. Environ. Res. Public Health*, 19, 2245. <https://doi.org/10.3390/ijerph19042245>.
- Yeni, S., & Cagiltay., K. (2017). A heuristic evaluation to support the instructional and enjoyment aspects of a math game. *Program*, 51(4), 406–423. <https://doi.org/10.1108/PROG-07-2016-0050>.
- Yilmaz, R. (2017). Exploring the role of e-learning readiness on student satisfaction and motivation in flipped classroom. *Computers in Human Behavior*, 70, 251-260. <https://doi.org/10.1016/j.chb.2016.12.085>.
- Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *Ieee Computational Intelligence Magazine*, 13(3), 55–75. <https://doi.org/10.1109/MCI.2018.2840738>.